Michael Wheeler

Thinking Beyond the Brain. Educating and Building from the Standpoint of Extended Cognition

2015

https://doi.org/10.25969/mediarep/1312

Veröffentlichungsversion / published version

Sammelbandbeitrag / collection article

Empfohlene Zitierung / Suggested Citation:
Wheeler, Michael: Thinking Beyond the Brain. Educating and Building from the Standpoint of Extended Cognition. In: Matteo Pasquinelli (Hg.): *Alleys of Your Mind. Augmented Intelligence and Its Traumas*. Lüneburg: meson press 2015, S. 85–104. DOI: https://doi.org/10.25969/mediarep/1312.

[5]

Thinking Beyond the Brain: Educating and Building from the Standpoint of Extended Cognition

Michael Wheeler

According to the hypothesis of extended cognition (ExC), our thinking is not just happening in the brain but spreads out to the beyond-the-skin environment. Following an introduction to the basic idea of extended cognition, this essay explores that idea in relation to two issues: first, it looks at the hybrid education in an increasingly networked world; second, at the situating of organic cognition within so-called "intelligent buildings." It is argued that we should understand these contemporary developments as the latest realizations of an age-old human ontology of dynamically assembled, organic-technological cognitive systems, since it is of our very nature to enhance our raw organic intelligence by forming shifting human-arte-fact coalitions that operate over various time-scales.

In *Alleys of Your Mind: Augmented Intellligence and Its Traumas,* edited by Matteo Pasquinelli, 85–104. Lüneburg: meson press, 2015.
DOI: 10.14619/014

We Have the Technology

In a widely reported article published recently in *Science* (Sparrow and Wegner 2011), a series of experimental results were described which together indicate that, in an era of laptops, tablets, and smartphones that come armed with powerful Internet search engines, our organic brains often tend to internally store not the information about a topic, but rather how to find that information using the available technology.

For example, in one experiment the participants were each instructed to type, into a computer, forty trivia statements that might ordinarily be found online (e.g., "An ostrich's eye is bigger than its brain"). Half the participants were told that their typed statements would be saved on the computer and half were told that their typed statements would be deleted. Within each of these groups, half of the individuals concerned were asked explicitly to try to remember the statements (where "remember" signals something like "store in your brains"). All the participants were then asked to write down as many of the statements as they could remember. The results were intriguing. The fact of whether or not a participant was asked to remember the target statements had no significant effect on later recall, but the steer about whether or not the statements would be saved on the computer did, with superior recall demonstrated by those participants who believed that their typed statements had been deleted. In other words, where the expectation is that information will be readily available via technology, people tend not to store that information internally. Further studies provided participants in the saved condition with additional information indicating where on the computer the saved statements were being stored (e.g., folder names). This scenario uncovered a more complex profile of organic memory allocation, suggesting that people don't internally store where to find externally stored items of information when they have internally stored the items themselves, but that they do internally store where to find externally stored items of information when they have not internally stored the items themselves. There is some evidence, then, that "when people expect information to remain continuously available (such as we expect with Internet access), we are more likely to remember where to find it than we are to remember the details of the item" (Sparrow and Wegner 2011).

Predictably, during the reporting of these experimental results, even the serious media couldn't resist engaging in some mild fear-mongering about the technology-driven degeneration of human intelligence. For instance, even though the British newspaper *The Guardian* published an article whose main text conveyed an accurate impression of the research in question, the piece invited some familiar contemporary anxieties, by virtue of its arguably sensationalist title, "Poor Memory? Blame Google" (Magill 2011). Such negative spin, it must be said, runs largely contrary to the experimenters' own interpretation of their results, in which one finds the more uplifting thought that what we have here is “an adaptive use of memory” in which “the computer and online search engines [should be counted] as an external memory system that can be accessed at will” (Sparrow and Wegner 2011, 3). Nevertheless, one can certainly see how the revealed pattern of remembering might be treated as evidence of some sort of reduction in overall cognitive achievement.

Thinking clearly about these sorts of issues requires (among other things, no doubt) a combination of historical perspective and philosophical precision concerning how we understand the technological embedding of our naked organic intelligence. The necessary historical perspective is nicely captured by Andy Clark’s memorable description of human beings as natural born cyborgs (Clarck 2003). What this phrase reminds us is that although it is tempting to think of our cognitive symbiosis with technology as being a consequence, as opposed to merely a feature of a world populated by clever computational kit, to do so would be to ignore the following fact: It is of our very nature as evolved and embodied cognitive creatures to create tools which support and enhance our raw organic intelligence by dovetailing with our brains and bodies to form shifting human-artefact coalitions operating over various time scales. This is no less true of our engagement with the abacus, the book, or the slide rule than it is of our engagement with the laptop, the tablet, or the smartphone. We are, and always have been, dynamically assembled organic-technological hybrids—systems in which a squishy brain routinely sits at the center of causal loops that incorporate not only non-neural bodily structures and movements, but also external, technological props and scaffolds: Technologies are, it seems, (part of) us.

The claim that technologies are (part of) us might seem like a metaphorical flourish—or worse, a desperate attempt at a sound-bite—but I mean it literally, and that’s where the philosophical precision comes in. We need to distinguish between two different views one might adopt hereabouts. According to the first, sometimes called the embodied-embedded account of mind, intelligent behavior is regularly, and sometimes necessarily, causally dependent on the bodily exploitation of certain external props or scaffolds. For example, many of us solve difficult multiplication problems through the exploitation of pen and paper. Here, a beyond-the-skin factor helps to transform a difficult cognitive problem into a set of simpler ones. Nevertheless, for the embodied-embedded theorist, even if it is true that one could not have solved the overall problem without using pen and paper, the pen-and-paper resource retains the status of an external aid to some internally located thinking system. It does not qualify as a proper part of the thinking system itself. Thus, the thinking itself remains a resolutely inner phenomenon, even though it is given a performance boost by its local technological ecology.[1]

The second view in this vicinity takes a more radical step. According to the extended cognition hypothesis (henceforth ExC), there are actual (in this world) cases of intelligent action in which thinking and thoughts (more precisely, the material vehicles that realize thinking and thoughts) are spatially distributed over brain, body, and world in such a way that the external (beyond-the-skin) factors concerned are rightly accorded cognitive status. Here, the term "cognitive status" tags whatever status it is that we ordinarily grant to the brain in mainstream scientific explanations of psychological phenomena. For the extended cognition theorist, then, the coupled combination of pen-and-paper resource, appropriate bodily manipulations, and in-the-head processing counts as a cognitive system in its own right, a system in which although the differently located elements make different causal contributions to the production of the observed intelligent activity, nevertheless each of those contributions enjoys a fully cognitive status. It is this more radical view that will concern us here.[2]

In the next section, I shall present an introduction to the basic shape of (one prominent form of) ExC. My primary aim in the paper as a whole, however, is not to explicate in detail or to argue for the truth of ExC. Rather, it is to explore ExC in relation to two socially charged issues that ask questions of us and about us in our contemporary human lives. Those issues are: first, how we should teach our children in an increasingly wired, wireless, and networked world (our opening example of strategic memory allocation will be relevant again here) and, second, how we should conceptualize our relationship with so-called intelligent architecture. Put more succinctly, I am going to say something about educating and building, from the standpoint of extended cognition.

The Functionalist Route to Extended Cognition

One of the things that has always struck me about ExC is the fact that although most philosophers and cognitive scientists tend to greet the view (at first anyway) with a mixture of consternation and skepticism, the possibility that it might be true is actually a straightforward consequence of what, despite the inevitable dissenting voices, probably still deserves to be called the house

1 The case for embodied-embedded cognition in its various forms has been made over and over again. For two philosophical treatments that stress the kind of interactive causal coupling just described see: Clark 1997; Wheeler 2005.

2 The canonical presentation of ExC is by Clark and Chalmers 1998. Clark's own recent defense of the view can be found in Clark 2008b. For a timely collection that places the original Clark and Chalmers paper alongside a range of developments, criticisms, and defenses, see Menary 2010.

philosophy in cognitive science, namely functionalism. In general terms, the cognitive-scientific functionalist holds that what matters when one is endeavoring to identify the specific contribution of a state or process *qua* cognitive is not the material constitution of that state or process, but rather the functional role which it plays in generating cognitive phenomena, by intervening causally between systemic inputs, systemic outputs, and other functionally identified, intrasystemic states and processes. Computational explanations of mental phenomena, as pursued in, say, most areas of cognitive psychology and artificial intelligence, are functionalist explanations in this sense.

A note for the philosophers out there: I have avoided depicting functionalism as a way of specifying the constitutive criteria that delineate the mental states that figure in our pre-theoretical commonsense psychology, e.g., as a way of specifying what it is for a person to be in pain, as we might ordinarily think of that phenomenon. This philosophical project, laudable as it was, has faced powerful criticisms over many years.[3] However, even if that particular functionalist project is now doomed to failure, the status of functionalist thinking within cognitive science remains largely unaffected. Good evidence for this resistance to contamination is provided by the fact that disciplines such as artificial intelligence and cognitive psychology have not ground to a halt in the light of the widely acknowledged difficulties with the traditional philosophical project. The underlying reason for the resistance, however, is that function-based scientific explanations of psychological phenomena—explanations which turn on the functional contributions of various material vehicles in physically realizing such phenomena—do not depend on giving functional definitions of those phenomena.[4]

What all this indicates is that if functionalism is true, then the hypothesis of extended cognition is certainly not conceptually confused, although of course it may still be empirically false. On just a little further reflection, however, it might seem that there must be something wrong with this claim, since historically the assumption has been that the cognitive economy of functionally identified states and processes that the functionalist takes to be a mind will be realized by the nervous system (or, in hypothetical cases of minded robots or aliens, whatever the counterpart of the nervous system inside the bodily boundaries of those cognitive agents turns out to be). In truth, however, there isn't anything in the letter of functionalism as a generic philosophical framework that mandates this exclusive focus on the inner (Wheeler 2010a; 2010b). After all, what the functionalist schema demands of us is that we specify the causal relations that exist between some target element and a certain set of systemic inputs, systemic outputs, and other functionally identified, intrasystemic elements. There is no essential requirement that the boundaries

3 For an introduction to the main lines of argument, see Levin 2010.
4 For a closely related point, see Chalmers 2008, foreword to Clark 2008.

of the system of interest must fall at the organic sensory-motor interface. In other words, in principle at least, functionalism straightforwardly allows for the existence of cognitive systems whose borders are located at least partly outside the skin, hence Clark's term "extended functionalism" (Clark 2008a; 2008b; see also Wheeler 2010a; 2010b; 2011a).

One pay-off from developing ExC in a functionalist register is that it gives the ExC theorist something she needs—assuming, that is, that she wants to call on one of the archetypal supporting arguments for the view, the argument from parity. Here is Clark's recent formulation of the so-called parity principle.

> If, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting as part of the cognitive process, then that part of the world is (for that time) part of the cognitive process. (Clark 2008b; drawing on Clark and Chalmers 1998)

As stated, the parity principle depends on the notion of multiple realizability: the idea that a single type of mental state or process may enjoy a range of different material instantiations. To see the connection, we need to be clear about how the parity principle works. It encourages us to imagine that exactly the same functional states and processes which are realized in the actual world by certain externally located physical elements are in fact also realized by certain internally located physical elements. Having done this, if we then judge that the internal realizing elements in question count as part of a genuinely cognitive system, we must conclude that so do the external realizing elements in the environment-involving, distributed case. After all, by hypothesis, nothing about the functional contribution of the target elements to intelligent behavior has changed. All that has been varied is the spatial location of those elements. And if someone were to claim that being shifted inside the head is alone sufficient to result in a transformation in status, from non-cognitive to cognitive, he would, it seems, be guilty of begging the question against ExC.

So that's how the parity principle works. Its dependence on multiple realizability becomes visible (Wheeler 2011a) once one notices that the all-important judgment of parity is based on the claim that it is possible for the very same cognitive state or process to be available in two different generic formats—one non-extended and one extended. Thus, in principle at least, that state or process must be realizable in either a purely organic medium or in one that involves an integrated combination of organic and non-organic structures. In other words, it must be multiply realizable. So, if we are to argue for cognitive extension by way of parity considerations, the idea that cognitive states and processes are multiply realizable must make sense. Now, one of the first things undergraduate students taking philosophy of mind classes are taught is that functionalism provides a conceptual platform for securing multiple realizability. Because a function is something that enjoys a particular kind of independence from its implementing material substrate, a function must, in principle, be multiply realizable, even if, in this world, only one kind of material realization happens to exist for that function.

Of course, even among the fans of ExC, not everyone is enamored by the parity principle (Menary 2007; Sutton 2010), and those who remain immune to its charms are often somewhat contemptuous of the functionalist route to ExC, but that's a domestic skirmish that can be left for another day. What cannot be ignored right now is the fact that neither the parity principle, nor functionalism, nor even the two of them combined, can carry the case for ExC. What is needed, additionally, is an account of which functional contributions count as cognitive contributions and which don't. After all, as the critics of ExC have often observed, there will undoubtedly be some functional differences between extended cognitive systems (if such things exist) and purely inner cognitive systems. So, faced with the task of deciding some putative case of parity, we will need to know which, if any, of those functional differences matter. In other words, we need to provide what Adams and Aizawa (2008) have dubbed a mark of the cognitive.

Even though I ultimately come out on the opposite side to Adams and Aizawa in the dispute over whether or not ExC is true, and even though (relatedly) I am inclined to dispute the precise mark of the cognitive that Adams and Aizawa advocate,[5] I do think we fundamentally agree on the broad philosophical shape that any plausible candidate for such a mark would need to take. A mark of the cognitive will be a scientifically informed account of what it is to be a proper part of a cognitive system that, so as not to beg any crucial questions, is fundamentally independent of where any candidate element happens to be spatially located (See Wheeler 2010a, 2010b, 2011a, 2011 b). Once such an account is given, further philosophical and empirical legwork will be required to find out where cognition (so conceived) falls—in the brain, in the non-neural body, in the environment, or, as ExC predicts will sometimes be the case, in a system that extends across all of these aspects of the world.

So that no one ends up feeling cheated, I should point out that nowhere in the present treatment do I specify in detail what the precise content of an ExC-supporting mark of the cognitive might be (see Wheeler 2011a). In relation to the present task of sketching functionalist-style ExC, I am interested only in the fact that the extended functionalist needs such a mark in order to determine which functional differences matter when making judgments about parity. That said, it is worth noting that the later arguments of this paper turn on a number of factors (including, for instance, functional and informational

5 A matter that I will not pursue here, but see Wheeler 2015.

integration, and a property that I shall call "dynamic reliability"), that are likely to feature when the necessary content is filled in.

The demand that any mark of the cognitive be scientifically informed reflects the point made earlier, that the functionalism that matters for ExC is the functionalism of cognitive science, not the functionalism that (some have argued—again, see above) characterizes commonsense psychology. In this context it is interesting to respond briefly to an argument from Clark to the effect that the fan of ExC should shun the idea of a mark of the cognitive (as I have characterized it) in favor of "our rough sense of what we might intuitively judge to belong to the domain of cognition" (Clark 2008b, 114). According to this view, judgments about whether or not some distributed behavior-shaping system counts as an extended cognitive system should be driven not by any scientific account of cognition, since such accounts are standardly "in the grip of a form of theoretically loaded neurocentrism" (Clark 2008b, 105), but rather by our everyday, essentially pre-scientific sense of what counts as cognitive, since the "folk [i.e., commonsense] grip on mind and mental state . . . is surprisingly liberal when it comes to just about everything concerning machinery, location, and architecture" (Clark 2008b, 106). Clark's claim strikes me as wrong (Wheeler 2011b). Indeed, there is good reason to think that the ordinary attributive practices of the folk presume the within-the-skin internality of cognition. Here is an example that makes the point. If an environmental protester had stolen the plans of Heathrow Terminal 5, in advance of the terminal being built, the folk would most likely have been interested, and either supportive of the act or outraged by it, depending on what other beliefs were in play. But presumably none of these attitudes would be held because the folk were considering the whereabouts of (to speak loosely) part of Richard Rogers' mind.[6]

We have now taken a brief stroll down the functionalist route to extended cognition and have highlighted (what I have argued are) three building blocks of that version of ExC—functionalism itself, the parity principle, and the mark of the cognitive. So, with ExC-functionalism style in better view, we can now turn our attention to those two aforementioned areas of contemporary life within which, I think, the notion of extended cognition has the potential to make itself felt, namely educating and building. My all-too-brief reflections on these issues are, of course, essentially those of the concerned citizen, since I am certainly no educational theorist and no architect. Like all philosophers, however, I feel I have the inalienable right to go wading around in other people's disciplines, although in my case I hope without any imperialistic tendencies. My humble goal is only to help initialize what hopefully turns out to be fruitful dialogues. So, with that goal in mind, let's begin with education.

6 Example taken from Wheeler 2011b.

Educating Extended Minds

Consider the following list of existing and potential examples of performance-enhancing technology that might be used in educational contexts: pen and paper; slide rules; limited capability generic calculators that have not been loaded with any personalized applications; restricted Internet access; largely unrestricted Internet access including the use of sophisticated search engines; the learners' own smartphones; sophisticated Internet search engines mainlined into the learners' brains via neural implants. (It might seem that the final example here is pure science fantasy, and maybe it is, but it is something that has at least been discussed hypothetically at Google. As Google's CEO Eric Schmidt mischievously reports in a 2009 interview: "Sergey [Brin] argues that the correct thing to do is to just connect [Google] straight to your brain. In other words, you know, wire it into your head.").[7] Given this list, we might echo some fears broached earlier, and ask ourselves the following question: assuming that, on average, overall behavioral performance will be better when the proficient use of technology is in place, does our list describe a slippery slope that marks the creeping degeneration of human intelligence or a progressive incline that shows our species the way to new cognitive heights?

One way of focusing the issue here is to ask under what conditions our children's intelligence should be formally examined, since, presumably, anyone who thinks that a cognitive reliance on increasingly sophisticated computational technology signals a degeneration of human intelligence will have a tendency not to want to see such technology readily allowed in examination halls. There is no doubt that, in some performance-testing contexts, we judge the use of performance-enhancing technology to be a kind of cheating. Sport provides obvious instances. Here is one illustrative case. Body-length swimsuits that improve stability and buoyancy, while reducing drag to a minimum, were outlawed by swimming's governing body FINA (Fédération Internationale de Natation) after the 2009 World Championships. In an earlier judgment that banned only some suits, but was later extended to include all body-length suits, FINA stated that it "[wished] to recall the main and core principle that swimming is a sport essentially based on the physical performance of the athlete."[8] One might try to export this sort of principle to our target case by arguing that "education is a process essentially based on the unaided cognitive performance of the learner," with "unaided" here understood as ruling

7 Michael Arrington, interview with Eric Schmidt, "Connect It Straight To Your Brain." Tech Crunch, 3 September 2009. http://techcrunch.com/2009/09/03/google-ceo-eric-schmidt-on-the-future-of-search-connect-it-straight-to-your-brain.

8 Quote retrieved from http://news.bbc.co.uk/sport1/hi/olympic_games/7944084.stm. Thanks to Andy Clark for suggesting this example to me.

out the exploitation of external technological resources.[9] On the basis of our exported principle, any technology that enhances the performance of the naked brain would be banned from the examination hall, although of course there would be no prohibition on the deployment of such technology as a kind of useful brain-training scaffold to be withdrawn ahead of the examination.

The foregoing reasoning is, of course, too simple in form. One complication is that we already partly test our children by way of research projects and other longer-term assignments that require the use of sophisticated computational technology, especially the Internet. Acknowledging this point, one might say that the question that concerns us at present is whether or not we should allow the same sort of technology to be used in all formal examinations. Here one might note that the combination of pen and paper already counts as a performance-enhancing technology that enables us to solve cognitive problems that our naked brains couldn't (see, for example, my earlier example of the way such technology figures in mathematical reasoning). Given the extra thought that the kind of contemporary technology that currently excites our interest is, in essence, just more of the performance-enhancing same (although of course much fancier in what it enables us to do), one might argue that we already have an affirmative answer to our question. The moot point, of course, is whether or not the path from pen and paper to smartphones and beyond is smoothly continuous or involves some important conceptual transition in relation to the matter at hand. In this context, another observation becomes relevant, namely that other examples of technology that appear earlier on (intuitively, at the less sophisticated end of) our list (e.g., generic calculators) are already allowed in examination halls, at least for certain tests. The fact that some technology is already deployed under examination conditions points to the existence of difficult issues about where on our list of performance-enhancing kit the transition from the permissible to the impermissible occurs, and about why that transition happens precisely where it does. As we shall see, such issues prompt further questions that receive interesting and controversial answers in the vicinity of ExC.

Many factors are no doubt potentially relevant to the kinds of issues just mentioned, some of which are not specific to the exploitation of the kind of external technology with which we are concerned. For example, I suspect (without, admittedly, having done any research beyond asking a few friends and colleagues) that many people (educationalists and the general public alike) would want to prohibit the use of some (hypothetical) genetically-tailored-to-the-individual synthetic cognitive booster pill taken just before an exam,

9 The case of neural implants that would enable mainline Google access is tricky to categorize, since such devices, although not of course the servers that they would access, would be located inside the cognizer's skin. To push on, let's just stipulate that neural implants count as external on the grounds that they are technological enhancements to organic intelligence.

but would want to allow the use of a performance-enhancing generic natural health supplement taken over many months, even if those two strategies had exactly the same outcome for the learner concerned (same grade, no ill effects on health, etc.). One thought that might be at work here (a thought that also seems to figure in questions of doping in sports) is that taking the long-term natural health supplement is, as its name suggests, a natural way of improving intellectual performance, whereas taking the immediate-effect tailored synthetic pill is an artificial prop. But whatever purchase this kind of thinking might have in the supplement-or-pill case, it seems questionable when we turn to the use of external technology such as search engines and smartphones, or at least it does if we view things from the standpoint of ExC. In actual fact, it already looks dubious from the less radical standpoint of embodied-embedded cognition, let alone ExC. That's because, according to both positions, human beings are (to recall once again Clark's phrase) natural born cyborgs. We have evolved to be (ExC), or to engage in (embodied-embedded view), shifting human-artefact coalitions operating over various time-scales. But if we really are natural born cyborgs, then the utilization of technology to enhance cognitive performance is as natural a feature of human existence as digestion or having children. So, on the suggested criterion, such utilization would fall on the permissible side of the divide.

It is possible, however, that the supplement-or-pill example introduces a different sort of consideration, namely whether or not the technology in question is generic (available in the same form to all, like the natural health supplement) or individualized (tailored to the individual, like the synthetic pill). Using this distinction as a way of cutting the cake, one might argue that generic technology (e.g., unrestricted Internet access via a shared search engine) is permissible in an exam setting, but individualized technology (e.g., the learner's own smartphone, loaded with personally organized information) is not. Once again, however, the truth of ExC would cast doubt on the proposed reasoning. One factor that will plausibly play a role in determining whether or not a particular external element is judged to be a proper part of an extended cognitive architecture is the functional and informational integration of that element with the other elements concerned, including of course those located in the brain. This integration will depend partly on the extent to which some external element is configured so as to interlock seamlessly with the desires, preferences and other personality traits that are realized within the rest of the cognitive system, a system which, of course, according to the ExC theorist, may itself be extended.

For example, compare a mobile application that recommends music to you purely on the basis of genre allocations with one whose recommendations are shaped by an evolving model not only of the kinds of purchases that you, as an individual, have made, but also of various psychological, emotional, political, and aesthetic patterns that your music-buying and other ongoing behavior instantiates. It seems that, if a suite of additional conditions were in place (e.g., real-time access of the applications when needed, a reliable pattern of largely uncritical dependence on the recommendations made), then the individualization demonstrated by the second program raises the chances that it deserves to be counted as part of your cognitive system (as partly realizing some of your beliefs and desires). But if that is right, then, from the standpoint of ExC, it is hard to see how the individual tailoring of an item of technology can be a sufficient reason to prohibit the use of that item in an examination. Such tailoring will, if other conditions are met, be part of an evidential package which (to employ what is, perhaps, an overly crude formulation) indicates that the technology in question counts as part of the learner's mind, and surely we want to allow that into the examination hall. From the standpoint of ExC, then, there seems to be no good reason based purely on individualization to ban sophisticated personal technology such as smartphones from any examination hall.

In response to this, someone might point out that our current examination rules, which sometimes allow certain items of technology (e.g., generic calculators) to be used in examination halls, are the result of context-dependent decisions regarding what it is that we are testing for. Thus, using a calculator might qualify as cheating in one sort of mathematics examination (in which we are testing for basic mathematical abilities), but be perfectly acceptable in another (in which we are testing for a more advanced application of mathematical reasoning). Although this might well be true, it seems, at first sight, that the ExC-driven reasoning that makes it acceptable to utilize those items of technology that achieve cognitive status, because they are dynamically integrated into the right sorts of causal loops, will enjoy a priority over any decisions based on the content of particular exams. After all, to replay the point made just a few sentences ago, from the standpoint of ExC, the technology in question has been incorporated into the learner's cognitive architecture (crudely, it is part of her mind), and that is the very "thing," it seems, that we are endeavoring to examine.

Once again, however, things are not quite so simple. This becomes clear once we recognize that the supporter of ExC will be driven to ask a slightly different question than "What are we testing for?" She will want to ask, "What are we testing?" To see why this is, recall the parity driven argument for ExC and the accompanying commitment to multiple realizability. These indicate that, for ExC as I have characterized it, the same type-identified psychological state or process, as specified functionally, will often be realizable in either a purely organic medium or in one that involves an integrated combination of organic and non-organic structures. So, nothing in ExC rules out the idea that cognition may sometimes be a wholly internal affair, which means that nothing in

ExC rules out the further idea that even though a person's cognitive system is sometimes extended, we might sometimes want to test the performance of her cognitive capacities under non-extended conditions. In other words, sometimes, we might still want to test the naked brain rather than the organic-technological hybrid. Where this is the case, we will want to ban the use of technology from the examination hall.

That said, one needs to be clear about what the motivation might be for testing the unadorned inner. After all, the experimental results described at the beginning of this paper indicate that when learners expect information to be readily and reliably available from an external resource (such as the Internet), they are more likely to remember where to find that information than the details of the information itself. This cognitive profile seems entirely appropriate for a world in which the skill of being able to find, in real time, the right networked information (not just facts, but information about how to solve problems) is arguably more important than being able to retain such information in one's organic memory. In such a world, which is our world, the brain emerges as a locus of adaptive plasticity, a control system for embodied skills and capacities that enable the real-time recruitment and divestment of technology in problem-solving scenarios. As such, and from the standpoint of ExC, the brain is most illuminatingly conceptualized as one element—albeit the core persisting element—in sequences of dynamically constructed and temporarily instantiated extended cognitive systems. Perhaps what we ought to focus on, then, is the education of those hybrid assemblages, a focus that is entirely consistent with the goal of endowing the brain with the skills it needs to be an effective contributor to such assemblages. From this perspective, of course, there are extremely good reasons to support the increased presence of technology in the examination hall. Moreover, it should be clear that, if ExC is right, then the list of technological entanglements within educational contexts with which we began this section reflects not the gradual demise of human intelligence in the age of clever computational kits, but rather our ongoing evolution as the organic-technological hybrids that we are, and that we have always been.

Dwellers on the Threshold

"I go up," said the elevator, "or down."
"Good," said Zaphod, "We're going up."
"Or down," the elevator reminded him.
"Yeah, OK, up please." There was a moment of silence.
"Down's very nice," suggested the elevator hopefully.
"Oh yeah?"
"Super."
"Good," said Zaphod, "Now will you take us up?"

> "May I ask you," inquired the elevator in its sweetest, most reasonable voice, "if you've considered all the possibilities that down might offer you?"

The preceding dialog is a conversation between Zaphod Beeblebrox and an elevator designed by the Sirius Cybernetics Corporation, from The Restaurant at the End of the Universe by Douglas Adams.[10]

Increasingly, architects will be designing buildings that, via embedded computational systems, are able to autonomously modify the spatial and cognitive environments of the people dwelling within them, in the light of what those buildings "believe" about the needs, goals, and desires of the people concerned. In other words, we are about to enter an era of intelligent architecture. Given our present concerns, the advent of such buildings invites the following question, for which I shall try to provide a preliminary answer: what is the relationship between ExC and the way in which we understand and conceptualize our cognitive relationships with intelligent buildings?

To focus our attention, let's get clearer about the intelligent architecture concept, and illustrate it with some examples. After a careful survey and analysis, Sherbini and Krawczyk (Sherbini and Krawczyk 2004, 150) define an intelligent building as one "that has the ability to respond (output) on time according to processed information that is measured and received from exterior and interior environments by multi-input information detectors and sources to achieve users' needs and with the ability to learn." Notice that Sherbini and Krawczyk's definition includes the requirement that the building should be able to learn, i.e., adjust its responses over time so as to provide the right environments for its users as and when those users need them. The idea that some sort of capacity to learn is a necessary condition for a building to be intelligent is one way of separating out the intelligent building concept from closely related notions, such as those of responsive architecture and kinetic architecture. The term "responsive architecture" applies to buildings that have the ability to respond to the needs of users. The term "kinetic architecture" applies to "buildings, or building components, with variable location or mobility, and/or variable geometry or movement" (Fox and Yeh 2011, 2). The variability involved in kinetic architecture may involve nothing more than opening a door or window, but it may involve moving a major structure which, in the limit, may be the whole building. The key thought behind the "separating out" move here is that not all responsive buildings, and not all kinetic buildings qualify as intelligent, since in some cases the responsiveness and/or the kinetic properties of those buildings will be the result of "unintelligent" processes such as direct, unmodifiable links between sensors and motors (cf. the idea that genuine intelligence in animals and humans requires more than

10 I have stolen the use of this quotation from Haque 2006.

hard-wired stimulus-response connections). Learning is one way to secure the right kind of "inner" mediation.

Against this conceptual backdrop, consider four examples of actual, planned, and exploratory buildings that are arrayed along a spectrum from mere responsive/kinetic architecture to intelligent architecture.

- Built in 1994, the Heliotrope, designed by Rolf Disch, is a kinetic building in Freiburg that, using solar trackers, rotates so as to follow the sun, thereby maximizing its access to solar energy and helping to minimize its heating energy demands from other sources. The Heliotrope was the first building in the world to generate more energy than it uses.[11]
- The Cybertecture Egg is a projected building, designed by James Law Cybertecture, to be located in Mumbai.[12] The building combines various intelligent, interactive, and multimedia systems to create an adapted and adaptable environment. Here are two examples: The bathrooms contain a system that monitors and records certain data indicative of the inhabitants' health (e.g., blood pressure, weight), data which may later be recovered and forwarded to a doctor; the inhabitants' working spaces may be customized to optimize individual experience (e.g., the actual view can be replaced by real-time virtual scenery retrieved from all over the world).
- Taking on the challenge of creating buildings in which the elderly can continue to live at home, the Ambient Assisted Living Research Department at the Fraunhofer Institute for Experimental Software Engineering in Kaiserslautern designed an intelligent embedded system that monitors the behavior of a building's inhabitants, via a network of hidden sensors (Kleinberger et al. 2009, 199–208). This network identifies and assesses risk situations (e.g., someone having a fall), and reports to a control center, allowing, say, the automatic notification of a designated contact. In addition, various intelligent systems autonomously modify the environment to reduce risk. Thus, the bathroom has a toilet that recognizes the user and adjusts itself to be at the appropriate height, and a mirror with illuminated pictograms that are designed to structure the activities of easily confused occupants by, for instance, guiding them to brush their teeth, wash, or take medication.
- In the exploratory architectural project Evolving Sonic Environment, developed by Haque and Davis (Haque 2006), people walk around inside an acoustically-coupled "spatialized" neural network (a spatial web of interconnected simple processing units). The movements of the occupants (detected via sound) affect the organization of the network (the architectural environment) through the operation of local learning algorithms

11 Rolf Disch, "Rotatable Solar House HELIOTROP: The experience of living rotating completely around the sun," architecture project, Freiburg, 1994. Published online: http://www.rolfdisch.de/files/pdf/RotatableSolarHouse.pdf.

12 See the projects section on the Cybertecture website: http://www.jameslawcybertecture.com

> active at each of its nodes. This results in the network adapting over time to different patterns of occupancy, often developing perceptual categories for reflecting those patterns that do not necessarily correspond to categories that the human observer would employ.

Now that we have intelligent architecture in view, we can investigate the relations between such architecture and ExC. Here is one way of asking the key question: Can the embedded systems in the walls and basements of intelligent buildings ever become constituent elements in the functionally specified material vehicles that realize the thoughts of those buildings' inhabitants? Put another way, could the sequence of dynamically assembled, organic-technological hybrid systems that instantiates my mind ever include factors embedded in the intelligent buildings in which I will increasingly dwell? To provide an answer here, I shall explore two lines of thought.

One factor that sometimes figures in discussions of ExC is the portability of cognitive resources. Indeed, it is sometimes suggested that a material element may count as the vehicle, or as part of the vehicle, of a thinker's cognitive state or process, only if that thinker carries, or at least is able to carry, the element in question around with her. In the language of section 2 (above), the portable-non-portable distinction marks a functional difference that matters when one is deciding whether or not a particular functional contribution to intelligent behavior counts as cognitive. Neural resources manifestly meet the proposed portability constraint. So too do PDAs and smartphones. Intelligent architecture, however, does not. So, if portability is a keystone requirement for a resource to be awarded cognitive status, then intelligent buildings are "no more than" adaptive scaffolds for richly coupled embodied-embedded minds, not vehicles for extended minds. But is portability what matters here? I don't think so. What really matters is a property in relation to which portability makes a positive enabling contribution, but which may be secured without portability. That property is somewhat difficult to specify precisely, but, roughly speaking, it amounts to a kind of dynamic reliability in which access to the externally located resource under consideration is, for the most part, smooth and stable just when, and for as long as, that resource is relevant to some aspect of our ongoing activity. The qualifier "dynamic" here reflects the fact that, according to ExC, the organism-centered hybrid systems that are assembled through the recruitment and divestment of technology often persist only when, and as long as, they are contextually relevant, meaning that the external resources concerned need not be smoothly and stably accessible at other times.

We can now state a modified condition for cognitive status: a material element may count as the vehicle, or as part of the vehicle, of a cognitive state or process, only if it meets the foregoing dynamic reliability constraint. And although carrying an item of technology around with you is certainly one assisting factor here, it is certainly not mandatory. Technological resources embedded in the fabric of one's house may well be readily and reliably available whenever the human behaviour that they support is operative. Consider, for example, the activity-structuring pictograms embedded in the mirrors of the ambient assisted living environment described earlier. When functioning in a hitch-free manner, access to these externally located resources will be smooth and stable just when, and for as long as, those resources are relevant to the activity they are designed to support. To be clear, meeting the dynamic reliability constraint in this way is clearly not a sufficient condition for a technological resource to count as part of one's cognitive architecture. But, if it is a necessary condition, then intelligent architecture may certainly, in principle, meet it.

Time, then, to turn to the second ExC-and-intelligent-architecture related issue that I want to broach here. Part of the interest of the final example of intelligent architecture described above, namely Evolving Sonic Environment by Haque and Davis, is that it foregrounds the already highlighted incorporation of learning into intelligent architecture. But the Haque and Davis study does more than that. It also introduces a new consideration, that of interaction. Haque argues that an important transformation in our relations with architecture occurs when we shift from a merely reactive kind of architecture to a genuinely interactive kind (Haque 2006).

Here Haque draws a distinction between single-loop interaction—in which the architectural response to a particular user-input is determined in advance—and multiple-loop interaction, in which the next response, by the architecture or user, is in part determined by an ongoing history of interaction and on the fact that each is able to access and modify each other's goals. As Haque puts it:

> [S]ingle-loop devices that satisfy our creature comforts are useful for functional goals (I am thinking here of Bill Gate's technologically-saturated mansion; or building management systems that seek to optimise sunlight distribution; or thermostats that regulate internal temperature). Such systems satisfy very particular efficiency criteria that are determined during, and limited by, the design process. However, if one wants occupants of a building to have the sensation of agency and of contributing to the organization of a building, then the most stimulating and potentially productive situation would be a [multi-loop] system in which people build up their spaces through "conversations" with the environment, where the history of interactions builds new possibilities for sharing goals and sharing outcomes. (Haque 2006, 3)

To put flesh (or perhaps concrete) on this goal of human-architecture conversation, Haque introduces his notion of Paskian Systems (named after the great maverick British cyberneticist, Gordon Pask). Paskian systems eschew the usual logic of the interaction between humans and smart technology. According to that usual logic, either the human user needs an appropriate understanding of the design of the machine, so that she can tell it what to do, or the machine needs an appropriate understanding of the design of the human user so that it can provide her with precisely what she needs. A Paskian system, by contrast, would support a kind of open dialog. Thus, for example, in a spatial dwelling context such a system "would provide us with a method for comparing our conception of spatial conditions with the designed machine's conception of the space" (Haque 2006, 3).

There is a compelling consideration which suggests that although the kind of non-Paskian architectural technology that we encountered earlier (recall, again, the mirror-embedded pictograms) may qualify as proper parts of the dweller's cognitive economy on roughly the same grounds as mobile computing technology (e.g., among other things, both meet the dynamic reliability constraint), Haque's Paskian systems—and thus the realizations of such systems in intelligent architecture—will fail to qualify. In fact, the threat to ExC here is established by the very conditions that make possible the capacity of Paskian systems to enter into richly interactive dialogs, the feature of those systems that secures Haque's advocacy of them in architectural design. Paskian systems may operate with categorizations, conceptions, and models of goal-states to be achieved—beliefs about how the dweller's world is and should be, if you will—that diverge from those of their human users. Thus, as mentioned earlier, the Evolving Sonic Environment develops perceptual categories for occupancy patterns that do not necessarily correspond to human-determined categories. It is this divergence that grounds the dialogical structure that characterizes the kind of rich human-building interaction sought by Haque. Now, this may well be exactly what we want from intelligent architecture, but the divergence calls into question any claim that the human-technology interactive system so instantiated is itself a single, integrated cognitive system. We would experience the same hesitation to think in terms of extended cognition if we were confronted by a Paskian smartphone that negotiated over where to go every time its online navigation program was fired up. And the same qualms indicate why the elevator designed by the Sirius Cybernetics Corporation (see above) cannot plausibly be considered part of Zaphod's mind.

The root issue here is that Paskian systems exhibit a kind of agency. This agency, however limited, prevents them from being incorporated into the cognitive systems that are centered on their human users. As one might put it, where there's more than one will, there's no way to cognitive extension. At first sight, this principle would seem to have negative implications (implications that I do not have the space to unravel or explore here) for the hypothesis of socially extended cognition, interpreted as the claim that some of the material vehicles that realize my thinking may be located inside the brains of other people (i.e., other agents). For the present, however, my thoughts are restricted to the domain of intelligent architecture: if intelligent architecture does support ExC, then it is on the basis not of Paskian interaction, but of the dynamic reliability established by non-Paskian loops.

Conclusion

The extended cognition hypothesis is currently the subject of much debate in philosophical and cognitive-scientific circles, but its implications stretch far beyond the metaphysics and science of minds. We have only just begun, it seems, to scratch the surface of the wider social and cultural ramifications of the view. If our minds are partly in our smartphones and even in our buildings, then that is not a transformation in human nature, but only the latest manifestation of the age-old human ontology of dynamically assembled, organic-technological cognitive systems. Nevertheless, once our self-understanding catches up with our hybrid nature, the world promises to be a very different place.

Acknowledgments: Thanks to Andy Clark for a discussion regarding education and cognitive extension that helped me to fine-tune my thinking on that issue. Thanks also to audiences at the Universities of Stirling and Sussex for useful feedback. This text has been previously published in the Open Access journal Computational Culture, no. 1, November 2011 (see: www.computationalculture.net).

References

Adams, Frederick, and Kenneth Aizawa. 2008. *The Bounds of Cognition*. Malden, MA and Oxford: Blackwell.

Chalmers, David J. 2008. "Foreword to Andy Clark's *Supersizing the Mind*." In Clark 2008b, ix–xvi.

Clark, Andy. 1997. *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.

Clark, Andy. 2003. *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. New York: Oxford University Press.

Clark, Andy. 2008a. "Pressing the Flesh: A Tension in the Study of the Embodied, Embedded Mind?" *Philosophy and Phenomenological Research* 76 (1): 37–59.

Clark, Andy. 2008b. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. New York: Oxford University Press.

Clark, Andy, and David J. Chalmers. 1998. "The Extended Mind." *Analysis* 58 (1): 7–19.

Fox, Michael A., and Bryant P. Yeh. 2011. "Intelligent Kinetic Systems." http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.124.4972&rep=rep1&type=pdf.

Haque, Usman. 2006. "Architecture, Interaction, Systems." Extended version of a paper written for Arquitetura & Urbanismo, AU149, Brazil, 2006. http://www.haque.co.uk/papers/ArchInterSys.pdf/, last accessed 13 October 2011.

Kleinberger, Thomas et al. 2009. "An Approach to and Evaluations of Assisted Living Systems Using Ambient Intelligence for Emergency Monitoring and Prevention." In *Universal Access* *in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, edited by Constantine Stephanidis, 199–208. Berlin: Springer.

Levin, Janet. 2010. "Functionalism". In *Stanford Encyclopedia of Philosophy*, Summer 2010 Edition. http://plato.stanford.edu/archives/sum2010/entries/functionalism.

Magill, Sasha. 2011. "Poor Memory? Blame Google." *The Guardian*, 16 July 2011. http://www.guardian.co.uk/science/2011/jul/15/poor-memory-blame-google.

Menary, Richard. 2007. *Cognitive Integration: Mind and Cognition Unbounded*. Basingstoke: Palgrave Macmillan.

Menary, Richard, ed. 2010. *The Extended Mind*. Cambridge, MA: MIT Press.

Sherbini, Khaled, and Robert Krawczyk. 2004. "Overview of Intelligent Architecture." Proceedings of the 1st ASCAAD International Conference, e-Design in Architecture, KFUPM, Dhahran, Saudi Arabia, December 2004. http://mypages.iit.edu/~krawczyk/ksascad04.pdf .

Sparrow, Betsy, Jenny Liu, Daniel M. Wegner. 2011. "Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips." *Science* 333 (6043), 776–78. http://www.sciencexpress.org .

Sutton, John. 2010. "Exograms and Interdisciplinarity: History, the Extended Mind, and the Civilizing Process." In: Menary 2010, 189–225.

Wheeler, Michael. 2005. *Reconstructing the Cognitive World: the Next Step*. Cambridge, MA: MIT Press.

Wheeler, Michael. 2010a. "In Defense of Extended Functionalism." In Menary 2010, 245–70.

Wheeler, Michael. 2010b. "Minds, Things, and Materiality." In *The Cognitive Life of Things: Recasting the Boundaries of the Mind*, edited by L. Malafouris and C. Renfrew. 29–37. Cambridge: McDonald Institute Monographs. Reprinted in *Action, Perception and the Brain: Adaptation and Cephalic Expression*, edited by Jay Schulkin, 147–63, Basingstoke: Palgrave Macmillan, 2012.

Wheeler, Michael. 2011a. "Embodied Cognition and the Extended Mind." In *The Continuum Companion to Philosophy of Mind*, edited by J. Garvey. 220–38, London: Continuum.

Wheeler, Michael. 2011b. "In Search of Clarity about Parity." *Philosophical Studies* (symposium on Andy Clark's *Supersizing the Mind*) 152 (3): 417–25.

Wheeler, Michael. 2015. "On Your Marks." Online Draft Manuscript, chapter 4. http://rms.stir.ac.uk/converis-stirling/publication/12884.