

Foreword to Andy Clark's *Supersizing the Mind*

David J. Chalmers

A month ago, I bought an iPhone. The iPhone has already taken over some of the central functions of my brain. It has replaced part of my memory, storing phone numbers and addresses that I once would have taxed my brain with. It harbors my desires: I call up a memo with the names of my favorite dishes when I need to order at a local restaurant. I use it to calculate, when I need to figure out bills and tips. It is a tremendous resource in an argument, with Google ever present to help settle disputes. I make plans with it, using its calendar to help determine what I can and can't do in the coming months. I even daydream on the iPhone, idly calling up words and images when my concentration slips.

Friends joke that I should get the iPhone implanted into my brain. But if Andy Clark is right, all this would do is speed up the processing, and free up my hands. The iPhone is part of my mind already.

Clark is a connoisseur of the myriad ways in which the mind relies on the world to get its work done. The first part of this marvelous book explores some of these ways: the extension of our bodies, the extension of our senses, and crucially, the use of language as a tool to extend our thought. The second part of the book defends the thesis that in at least some of these cases, the world is not serving as a mere instrument for the mind. Rather, the relevant parts of the world have become parts of my mind. My iPhone is not my tool, or at least it is not wholly my tool. Parts of it have become parts of me.

This is the thesis of the extended mind: when parts of the environment are coupled to the brain in the right way, they become parts of the mind. The thesis has a long history: I am told that there are hints of it in Dewey, Heidegger, and Wittgenstein. But no-one has done as much to give life to the idea as Andy Clark. In a series of important books and articles—*Being There*, *Natural-Born Cyborgs*, “Magic words: How language augments human computation”, and many others—he has explored the many ways in which the boundaries between mind and world are far more flexible than one might have thought. This book is his major statement of the philosophical picture that

undergirds the view.

Andy invited me to write this foreword because of my role in co-authoring an article, “The Extended Mind” (included as an appendix to this book), that has come to serve as a sort of flagship philosophical statement for this picture of the mind. This paper was written when Andy and I were colleagues at Washington University in 1995, and was published in *Analysis* in 1998. Now that a decade has passed, Andy suggested that I might offer a retrospective perspective on that article while saying something about my own views on the topic. I am more than happy to do this, though no-one should feel any obligation to read what I have to say. Readers who are not already familiar with the issues might want to first look at the rest of the book, or at least at the appendix, before returning to this foreword.

It will come as no surprise to anyone who knows Andy’s work to hear that the inspiration behind the original article was all his. In March 1995, Andy handed me a short article that he had written, called “Mind and World: Breaching the Plastic Frontier”. This article already contained many of the key theses and arguments that appear in “The Extended Mind”. It contained the key thought-experiment comparing Tetris players who rotate images inside and outside the head. It also contained the crucial Parity Principle, which held that if a process in the world works in a way that we should count as a cognitive process if it were done in the head, then we should count it as a cognitive process all the same. I had some ideas about how to further develop and argue for the thesis, and we ended up working together on an expanded and renamed paper.

The original article contained a notorious footnote, saying “The authors are listed in order of their degree of belief in the central thesis”. Some have taken this to imply that I reject the extended mind thesis, and was working as a sort of hired gun in service of someone else’s cause. In fact, I find the thesis very attractive. If anything, I am more confident of the thesis than I was a decade ago, having seen the thesis survive the many objections that have been raised to it without too much trouble. (I agree with almost everything in Andy’s authoritative treatment of these objections in part two of this book.) But I do not think the matter is completely cut and dried.

I am not worried by the most common objections to the thesis: the fact that external cognitive processes work differently from internal cognitive processes, the threat that the mind will extend too far into the world, and the threat that the core role of the brain will be lost, for example. And I certainly do not think that there is anything privileged about skin and skull as boundaries for the mind.

Still, I think that there is one potentially principled place where the opponent of the extended mind can resist. This is an appeal to the dual boundaries of perception and action. It is natural

to hold that perception is the interface where the world affects the mind, and that action is the interface where the mind affects the world. If so, it is tempting to hold that what precedes perception and what follows action is not truly mental. And one might use this to draw a principled distinction between the cases of Otto (the Alzheimer's patient who uses a notebook as memory) and Inga (the ordinary subject who uses her brain). To interact with his notebook Otto must read it and write in it, requiring perception and action, where there is no such requirement for Inga. If so, then the boundaries above would place the notebook outside the mind.

We consider this sort of worry briefly in "The Extended Mind", suggesting that Otto's access to the notebook need not be seen as perceptual. But this is surely too quick: there is no denying that Otto sees the notebook and reads from it, just as there is no denying that Otto reaches for the notebook and writes in it. So there is certainly perception and action taking place here. A better reply might be to note that there can be inner perception (when one reads from a mental image, say) and mental action (when one makes a mental note, say). Then Otto's perception and action could be seen as of a piece with these. But an opponent could reasonably respond that Otto's interaction with the notebook involves real perception and real action (where these might be glossed as sensory perception and physical action, or in terms of the right sort of perceptual and agentic experience, or in some other way), of a sort not present in cases of inner perception and mental action. And it could be suggested that real perception and real action mark a plausible boundary for the mental.

Perhaps the best reply for the proponent of the extended mind is to reject the proposed boundary. We touch on this reply in the article, saying that just because the Terminator retrieves information by reading it from a screen, this does not mean that the information is not truly part of its memory. One can more generally hold that the difference between real perception and inner perception, or between real action and mental action, is not sufficiently robust or important to ground a mental/non-mental distinction. Still, an opponent could hold the line, saying that if the Terminator has to retrieve the information by reading it, then he did not truly believe it beforehand. And perhaps this opponent would have commonsense psychology on their side. If so, then perhaps this is one point where the "commonsense functionalism" that Clark favors in this book, individuating mental states by the roles that commonsense psychology assigns to them, counts against the extended mind thesis.

At this point, I think the proponent of the extended mind should not be afraid of a little revisionism. Even if commonsense psychology marks a distinction here, the question still arises of whether this is an important distinction that ought to be marked in this way. One can argue, as

we do in the article, that Otto's extended state involving the notebook functions in explanation in very much the way that beliefs function in psychological explanation. If so, then it ought to be classified as a belief, whether or not it is so classified by common sense. And it is a familiar philosophical move to argue that if a state shares the most important explanatory features of a belief, then it really is a belief.

Still, this strategy will work only if the involvement of perception and action makes no important difference to the explanatory role of Otto's extended state. And this is not so clear. For at least some explanatory purposes, these things seem to make a difference. After all, one can ask a crucial question: why did Otto reach for his notebook? This seems to be a perfectly good psychological question about the explanation of action. And the natural answer is: he wanted to get to the museum, he did not know its location, and he believed that the notebook contained the information. In this explanatory structure, we speak naturally as if Otto lacked the extended belief. On the other hand, for many other explanatory purposes, the mediating role of perception and action seems quite unimportant. We can ask, why did Otto walk north, and explain this in terms of his extended standing belief about the museum's location, in just the same way that we explain Inga's movements in terms of her beliefs.

I take the moral here to be that the classification of states can depend on our explanatory purposes. When we are interested in explaining Otto's large-scale behaviour, it is natural to say that his beliefs are extended, with his interactions with his notebook serving as a sort of uninteresting background constant. When we are interested in explaining Otto's local-scale interactions with his notebook, it is natural to deny that he has the extended belief, and to hold that the relevant actions are explained by internal beliefs. As Clark suggests in his concluding chapter, we can flip back and forth between both ways of looking at things. We have a sort of Necker Cube effect, with mental states counting as extended or not depending on our perspective and our purposes.

As this Necker Cube flips, various things flip along with it. Otto's access to the notebook flips from an act of perception to an act of memory retrieval. His writing in the notebook flips from a physical action to a mental action. We might thus flip from regarding Otto's cognitive system as local to regarding it as extended, and we might even flip our perspective on Otto himself in a similar way. And importantly, Otto's state before opening the notebook flips from ignorance to knowledge

This duality of perspectives can naturally be accommodated in various stories about the way mental terms like 'believe' and 'know' function. One such story holds that that ascriptions such as 'Otto believes that the Museum is on 53rd Street' are sensitive to contextual factors including

explanatory purposes. In a context where one is explaining Otto's travels, the ascription is true. In a context where one is explaining his interactions with his notebook, the ascription is false. One can find a similar dependence on explanatory purposes in non-extended cases. Say that someone has to think for a moment before responding that 7 times 8 is 56. If we are asked, why did they hesitate, we might reasonably say that they did not know the answer and had to think about it. But if we are asked, why did they get all the questions right, we might say that it is because they knew all the values in the multiplication table.

Other semantic stories are possible. One could give an account where the extended belief and knowledge ascriptions are always true, and when we say that Otto reached for his notebook because he did not know the address, we say something that is useful but strictly speaking false. One could also give an account where the extended belief and knowledge ascriptions are always false, so that our explanation of Otto's actions in terms of extended beliefs are at best metaphorically true. Or one could suggest that the reference of terms like 'belief' is indeterminate between the two notions. If so, then the original extended belief ascriptions may be neither true nor false, but there will be precisified versions that fall on each side.

Ultimately, however, I think that issues about what really counts as a belief and about how the term 'belief' functions are terminological questions that while interesting, can mask the deeper point. If someone insists that they use the term 'believe' in such a way that it picks out states realized in the space between perception and action, then one can allow them to use the term this way if they like. The deeper point is that extended states can function in explanation in very much the same way that beliefs function, and should be regarded as sharing a deep and important explanatory kind with them. This explanatory unification is the real underlying point of the extended mind thesis.

In "The Extended Mind", the only extended elements of the mind that we argued for were beliefs and cognitive processes: in particular standing beliefs (like Inga's belief about the museum's location), and cognitive processes such as mental rotation. It is natural to ask whether the extended mind thesis might itself be extended. What about extended desires, extended reasoning, extended perception, extended imagination, and extended emotions? I think that there is something to be said for each of these. Perhaps the camera on my iPhone can serve as an extended perceptual mechanism. And perhaps one might have something akin to an extended mood, if not an extended emotion, when one's environment is always nudging one toward happiness or sadness. Clark discusses many such cases throughout the book, including extended perceptual mechanisms in Chapter 2 and extended mechanisms of attention in Chapter 3.

But then, what about the big question: extended consciousness? The dispositional beliefs, cognitive processes, perceptual mechanisms, and moods considered above all extend beyond the borders of consciousness, and it is plausible that it is precisely the non-conscious part of them that is extended. I think there is no principled reason why the physical basis of consciousness could not be extended in a similar way. It is probably so extended in some possible worlds: one could imagine that some of the neural correlates of consciousness are replaced by a module on one's belt, for example. There may even be worlds where what is perceived in the environment is itself a direct element of consciousness: my paper "Perception and the Fall from Eden" tells a fable about one such world.

Still, I think it is unlikely that any everyday process akin to Otto's interaction with his notebook will yield extended consciousness, at least in our world. Certainly, relatives of the Otto/Inga argument do not seem to extend to consciousness. The original argument crucially yields a twin case, involving Otto and Twin Otto, who are physical duplicates with different beliefs. An argument for extended consciousness would require twins with different states of consciousness: Olga and Twin Olga are internal duplicates, but what it is like to be Olga differs from what it is like to be Twin Olga. But no matter how hard one tries to construct an Otto-style story that works like this, the story does not seem to succeed. Perhaps part of the reason is that the physical basis of consciousness requires direct access to information on an extremely high bandwidth. Perhaps some future extended system, with high-bandwidth sensitivity to environmental information, might be able to do the job. But our low-bandwidth conscious connection to the environment seems to have the wrong form as it stands.

In recent years, a few philosophers have argued that the basis of conscious states lies partly outside the head. Some of these arguments, such as those of Dretske (1996), Fisher (2007), and Martin (2004), turn on considerations quite different from the sort of two-way coupling between organism and environment that is at the heart of the extended mind thesis. The resulting views are interesting and challenging, but they are largely independent of active externalism of the extended mind thesis. Others, including Hurley (1998) and Noe (2006), have argued that the two-way coupling extends to consciousness. But these arguments do not seem to yield a twin case of the sort discussed above, so they do not rule out the supervenience of consciousness on the internal. At best, as Clark suggests in Chapter 8, they yield a weaker sort of dependence of consciousness on the environment. I tentatively conclude that the extension of the mind is compatible with retaining an internal conscious core.

What general picture of the mind does the extended mind thesis rest upon? It has sometimes

been suggested that the thesis requires functionalism about the mental, where all mental states are defined by the causal roles that they play. This cannot be quite right: I think that functionalism about consciousness is implausible, for example, but this implausibility does not affect the arguments for the extended mind thesis. One might support the view by invoking an attenuated functionalism: say, one where certain mental states (such as dispositional beliefs) are defined by their causal relations to conscious states, to behavior, and to other elements of the cognitive network. I find such a picture attractive myself, but strictly speaking even this picture is not required for all the argument to go through. All one needs is the very weak functionalism captured in the parity principle: roughly, if a state plays the same causal role in the cognitive network as a mental state, then there is a presumption of mentality, one that can only be defeated by displaying a relevant difference between the two (and not merely the brute difference between inner and outer). Combined with the observation that there are no relevant differences in the relevant cases—an observation that does not require functionalism for its support—the thesis follows.

Likewise, the extended mind thesis is compatible with both physicalism and dualism about the mental. It is compatible with connectionist and classical views, with computational and noncomputational approaches, and even with internalism and externalism in the traditional debates over mental content (as we suggest in a footnote to the paper). So I do not think that the extended mind thesis requires much in the way of theoretical presupposition at all. Instead, it is an independently attractive view of the mental.

Ultimately, the proof is in the pudding. The deepest support for the view comes from the explanatory insights that the extended mind perspective yields. And those insights are just what this book provides. In case after case, in domain after domain, Andy Clark bring out the many ways in which the extended view of the mind can productively reconfigure our thinking about the relationship between mind and world. After absorbing this picture, nothing will ever look quite the same way again. And if Clark is right, then the absorption has already started. Just opening his book may have turned you into a smarter, deeper, and more insightful person.

David Chalmers

Canberra, November 2007.

References

Chalmers, David J. (2006). Perception and the fall from Eden. In Tamar S. Gendler & John Hawthorne (eds.), *Perceptual Experience*. Oxford University Press.

Dretske, Fred (1996). Phenomenal externalism, or if meanings ain't in the head, where are qualia? *Philosophical Issues* 7.

Fisher, Justin C. (2007). Why nothing mental is just in the head. *Nous* 41:318-334.

Hurley, Susan L. (1998). Active perception and vehicle externalism. In Susan L. Hurley (ed.), *Consciousness in Action*. Harvard University Press.

Martin, M. G. F. (2004). The limits of self-awareness. *Philosophical Studies* 120:37-89.

Noe, Alva (2006). Experience without the head. In Tamar S. Gendler & John Hawthorne (eds.), *Perceptual Experience*. Oxford University Press.