

Karina Vold

## OVERCOMING DEADLOCK: SCIENTIFIC AND ETHICAL REASONS TO EMBRACE THE EXTENDED MIND THESIS

### ABSTRACT

The extended mind thesis maintains that while minds may be centrally located in one's brain-and-body, they are sometimes partly constituted by tools in our environment. Critics argue that we have no reason to move from the claim that cognition is *embedded* in the environment to the stronger claim that cognition can be *constituted* by the environment. I will argue that there are normative reasons, both scientific and ethical, for preferring the extended account of the mind to the rival embedded account.

### KEYWORDS

extended cognition, intracranialism, embedded mind, ethics, theory selection, cognitive rehabilitation

### 1. Introduction

Andy Clark and David Chalmers's extended mind thesis maintains that while minds are centrally located in one's brain-and-body they are sometimes partly constituted by tools in our environment. Some critics argue that we have no reason to move from the claim that cognition is *embedded* in the environment to the stronger claim that cognition can actually be *constituted* by the environment. In this paper I argue there are normative reasons, both scientific and ethical, for preferring the extended view to the embedded view. In the first place I appeal to the scientific values of simplicity, usefulness, and explanatory power, to argue that our best scientific theory of the mind will include extended mental states. In the second place, reviewing the literature, I appeal to three ethical reasons to prefer the extended view: it better protects against harm to the mind Levy (2007a,b); it better accounts for compensatory rehabilitation as a way of repairing the mind (Drayson and Clark, forthcoming); and it offers a better assessment of the capacities of learning disabled individuals (King 2016). All of these reasons, especially taken together, I argue provide some support for preferring the extended mind thesis to the rival embedded mind thesis.

### 2. Clark and Chalmers's Extended Mind Thesis

The dominant view in the brain sciences is that the brain alone constitutes, or realizes, the mind. But several philosophers have argued that the mind is sometimes

---

Karina Vold: Research Associate Ph.D, Leverhulme Centre for the Future of Intelligence, Faculty of Philosophy, University of Cambridge; kvv22@cam.ac.uk.

This work was supported by the Leverhulme Centre for the Future of Intelligence, Leverhulme Trust, under Grant RC-2015-067.

partially constituted by more than just the brain. While this idea has long roots in philosophy (Martin Heidegger, Maurice Merleau-Ponty, and John Dewey have all defended versions of this claim, for example), various contemporary philosophers have given new life to the idea, often under the title of *cognitive extension* or the *extended mind thesis*. In this paper I will focus on one particular version of this claim: the “extended mind” thesis as defended by Andy Clark and David Chalmers (1998). Clark explains, “[p]roponents of the extended mind story hold that even quite familiar human mental states (e.g., states of believing that so and so) can be realized, in part, by structures and processes located outside the human head.” (Clark 2008: 76) In other words, the extended mind thesis maintains that mental states (and processes) sometimes *extend* beyond the brain in the sense that they are partially constituted by extra-bodily states (or processes) working together with brain states (or processes).

To defend the extended mind thesis Clark and Chalmers argue,

[i]f, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* (so we claim) part of the cognitive process. (Clark and Chalmers 1998: 8)

This is now often referred to as their ‘parity principle’, while others call it the ‘fair treatment principle’ (e.g. Sprevak 2009; Drayson 2010) as it maintains that we should regard equivalent processes in similar ways, irrespective of whether they are internal or external to the skull. Motivated by this principle, their parity argument can be summarized:

- (P1) A physical state (or content-bearing structure) *p* is constitutive of a mental state of type *m* when *p* plays the causal role characteristic of *m* in the system.
- (P2) A physical state (or content-bearing structure) *p* located beyond (or partially beyond) an agent’s biological body can play the same causal role as physical states of the biological body that surely constitute an ordinary mental state of type *m*.
- (C) Therefore, physical states (or content-bearing structures) located beyond (or partially beyond) the biological body can be constitutive of an agent’s mental state.

The first premise expresses a commitment to a coarse-grained common-sense role functionalism.<sup>1</sup> To support the second premise, Clark and Chalmers describe a case in which, they argue, an object in the environment does play exactly the same role for one agent that neurons in the brain (something we would surely count as part of the supervenience or realization base of the mind) do for another. The case involves two people: Inga and Otto. Inga decides to go to an exhibition at the museum and to do so, “[s]he thinks for a moment and recalls that the museum is on 53rd Street, so she walks to 53rd Street and goes into the museum.” (Clark and Chalmers 1998: 12) Meanwhile, we imagine that Otto suffers from Alzheimer’s disease and has to rely on information he stores in a notebook to help structure his life. When he decides to go to the same exhibition he consults his notebook, where he

1 For the sake of brevity I will bypass an explanation of what this view maintains exactly. It is a popular, though not uncontroversial, view in contemporary analytic philosophy of mind.

has written the address and directions for how to get there. He then walks to the museum and heads inside (Clark and Chalmers 1998: 12–13). Clark and Chalmers argue that in the “relevant respects” the information in Otto’s notebook “functions just like” the information in Inga’s brain that constitutes an ordinary belief and thus both should count equally as part of the constitutive machinery of his mind (Clark and Chalmers 1998: 13). In other words, the information stored in the notebook *really is* a part of Otto’s mind, just like the information stored in Inga’s brain *really is* a part of her mind. More specifically, the information in the notebook is meant to be an example of an extended standing (i.e. not currently being entertained), non-conscious belief.

### 3. The Coupling-Constitution Objection

Importantly, the claim made by extended mind theorists is a *constitutive* one – that mental states and processes can be partially *constituted* by objects located beyond the brain and body. One of the major objections confronting the extended mind thesis maintains that we have no reason to move from the claim that cognition is *causally reliant* on the environment to the claim that cognition is partially *constituted* by the environment. Frederick Adams and Kenneth Aizawa argue that Clark and Chalmers commit a coupling-constitution fallacy by mistaking the mere causal dependence, or *coupling*, of extra-neural resources with neural activity for their *constitutive* involvement in unconscious mental states (or processes), such as Otto’s belief about the location of the museum.<sup>2</sup> Adams and Aizawa maintain that the *mere* coupling of a resource to a system does not imply the resource is partially constitutive of that system. The circulatory system is coupled to the cognitive system in the sense that circulation *causally supports* cognition in a crucial way, but this does not imply that circulation is partially constitutive of cognition (Adams and Aizawa, 2008: 10–11). Thus, to say that an object, x, is coupled to another, y, does not imply that x constitutes (or is a part of y). So to show that external objects are coupled to our cognitive states or processes does not imply that these partially constitute our cognition.

#### 3.1 The Embedded Mind Thesis

The distinction between the mere coupling of inner and outer resources and the constitutive involvement of outer resources is what distinguishes the *embedded* theory of cognition from the stronger *extended* mind theory. The embedded account maintains that a cognitive system depends, sometimes *crucially*, on the complexity of its environment, but that the environment is not an actual part of the mind. Herbert Simon, for example, argues that much of the apparent complexity of cognitive systems is actually external to the agent, residing in the environment. On this view cognitive systems lean heavily on this worldly complexity without internalizing it (Simon 1969: 51–52). For example humans sometimes structure their own environment to store information and then rely on these external structures instead of

2 See Adams and Aizawa (2001), (2008).

relying on internal resources. The *mise en place* method of lining up one's ingredients in the correct order for cooking, for instance, is widely used by chefs to save them from having to remember the ordering of their recipes while cooking (Clark 2008). The embedded view of cognition tells us that in order to understand and explain cognitive processes, such as the chef's use and processing of information while cooking, cognitive science cannot just study the internal processes of computation instantiated in the brain. Instead, we must study the way that structures in the local environment of an agent facilitate the success of the agent's internal processes. Thus, the embedded view offers an explanatory, or epistemic, reason to look beyond the brain. But the embedded view does not make any substantial constitutive claim; it does not challenge the (metaphysical) view that the brain wholly constitutes the mind (see Rowlands 2010, Chapter three for further discussion). Let us call this position, i.e. the view that the brain wholly constitutes the mind, 'intracranialism'. The key difference then is that embedded mind theorists accept intracranialism, while extended mind theorists reject it.

Extended mind theorists tend to think that embedded claims risk triviality. Almost everyone agrees that the mind is in some sense causally reliant on the body and the extra-bodily world. Many even agree that the body and tools in our environment can work as scaffolds, causally contributing to the development of certain higher-level cognitive capacities or the execution of cognitive tasks, such that these capacities or tasks would not have developed or could not be executed (or may not even confront the agent in the first place) were it not for these extra-cranial contributions. This view about 'cognitive scaffolding' was advocated by the developmental psychologist Lev Vygotsky (1978). But scaffolding falls short of the constitutive claim that the extended mind thesis makes (for more on this issue see Adams and Aizawa 2001, 2008; Shapiro 2008; Clark 2008, 2010).

Those who prefer the embedded mind theory, on the other hand, tend to think that there is no good reason for preferring the stronger constitutive claim made by the extended mind thesis to the more conservative coupling claim. In response to Adams and Aizawa's coupling-constitution objection, Clark (2010) argues that the burden of proof lies with those who reject the constitutive claim. They need a principled reason for maintaining that all mental states are entirely constituted by neural resources and only causally supported by extra-neural ones. To this end that Adams and Aizawa (2005) argue that *original*, or non-derived, content is the distinguishing 'mark of the cognitive. *Original*, or *non-derived*, content is meant to contrast with the *derived* content that non-mental objects can display. So, while non-mental objects, such as the words in a book, e.g. Otto's notebook, may carry *derived* content, original content is uniquely a feature of minds.<sup>3</sup> Thus, Adams and Aizawa argue that the notebook fails to be partially constitutive of Otto's mind because it lacks this 'mark of the mental' that — just as a matter of current contingent fact, on their view — only brains can realize.<sup>4</sup>

3 This distinction seems to be inspired by a distinction between derived and original intentionality made by Searle (1992).

4 I will not respond to Adams and Aizawa's objection here, though I believe there are several responses one could give (for some responses see Clark 2008b, 2010). Notice that even if the objection works against the example of Otto and his notebook, their distinction

But some who reject the extended mind theory have also rejected Clark's insistence that the burden of proof lies with them. What reasons do we have for preferring the more radical idea that technology, such as pens and papers and even smart phones, can really be partly constitutive of our mental life? Rupert (2004), for example, argues that there is no obvious reason for preferring the extended account. He argues, first, that adopting the embedded account is enough to recognize the indispensability of studying an agent's environment for understanding his cognition without conceding that the environment is actually partially constitutive of cognition. Furthermore, Rupert argues that we can explain all of the relevant phenomena that cognitive scientists study with the embedded account and, thus, moving to the stronger extended account is unjustified and unnecessary (Rupert 2004: 8–9). If the extended account does not offer us anything more than the embedded account, then following the methodological principle of conservatism, Rupert argues, we should endorse the embedded view over the extended mind view (Rupert 2004: 9). In this case, those defending the more radical extended view need to justify their proposed revision of our common-sense views about where the mind is.<sup>5</sup> Since this debate took off, several normative reasons for preferring the extended account over the rival embedded account have been suggested. I will now survey and evaluate these reasons as well as offer my own.

## 4. Scientific Virtues in Theory Selection

Thomas Kuhn (1977: 321–323) identifies several virtues or preferable characteristics that provide the basis for choice between competing scientific theories. These include accuracy, consistency, scope, simplicity, and fruitfulness. Kuhn thinks that accuracy is the most important of these virtues, but we've seen that there is currently a stalemate in the debate between which view—embedded or extended—is accurate. Thus, I will make the case that by appealing to other scientific values, including simplicity, usefulness and explanatory power, we are compelled to prefer the extended account to the embedded account. While not identified by Kuhn, usefulness and explanatory power are arguably accepted scientific virtues (especially the latter), and should be distinguished from Kuhn's other virtues, such as simplicity and fruitfulness, as they will sometimes be traded off against them.

### 4.1 Simplicity in Theory Selection

Simplicity is widely accepted as a norm of theory formulation in a wide range of disciplines, both humanistic and scientific. Occam's razor, for example, has long

---

would not block all possible cases of extension. We can imagine that instead of using a notebook Otto uses the mind of another agent, e.g. his long-time partner, to store the information that forms his beliefs. In this case the external resource is itself a brain, so it would be capable of original content. The result would be an instance of a *socially extended mind* — where one agent's mind has extended into another's brain ('social' because two agents are involved).

<sup>5</sup> Especially since Clark and Chalmers' argument appeals to a version of 'common-sense' functionalism.

been employed by philosophers as a way of guiding our preferences when choosing between two competing hypotheses. This principle maintains that, in the case where all other things are equal, for example, where both hypotheses can account equally as well for the data, we ought to opt for the simplest hypothesis, that is, the one that posits the fewest metaphysical entities. With respect to competing theories about where the supervenience base of the mind is, Clark and Chalmers (1998) suggest that we can use simplicity as a way of assessing and arbitrating between the extended view and the embedded view.

One reason folk psychology has endured is, quite plausibly, its simplicity (it of course also has other virtues, including explanatory power, which I discuss below). The belief-desire-intention (BDI) model of our practical reasoning, developed by Michael Bratman (1987), is a way of explaining how we perform everyday actions. It is also an example of the method we commonly use in explaining the actions of others in everyday life, which relies, fundamentally, on folk-psychological concepts. Clark and Chalmers use this model to argue that the extended account yields the simplest explanation of Otto's action:

Certainly insofar as beliefs and desires are characterized by their explanatory roles, Otto's and Inga's cases seem to be on a par: the essential causal dynamics of the two cases mirror each other precisely. We are happy to explain Inga's action in terms of her occurrent desire to go to the museum and her standing belief that the museum is on 53rd street, and we should be happy to explain Otto's action in the same way. (Clark and Chalmers 1998: 13)

In fact, according to the BDI model, we would explain Inga's action in terms of her occurrent desire to go to the museum, her standing belief about where the museum is located *and* her intention to take action. The extended mind thesis allows for the simplest application of the BDI model to Otto: we would explain Otto's action in terms of his occurrent desire to go to the museum, his standing belief about where the museum is located—which happens to be stored in his notebook, instead of in his brain—and his intention to take action. Thus, the extended view of the mind allows for the simplest application of the BDI model, a method of folk psychology that we regularly use. Consider the alternative explanation, as Clark and Chalmers describe it:

The alternative is to explain Otto's action in terms of his occurrent desire to go to the museum, his standing belief that the Museum is at the location written in the notebook, and the accessible fact that the notebook says the Museum is on 53rd Street; but this complicates the explanation unnecessarily. If we must resort to explaining Otto's action this way, then we must also do so for the countless other actions in which his notebook is involved; in each of the explanations, there will be an extra term involving the notebook. We submit that to explain things this way is to take one step too many. It is pointlessly complex, in the same way that it would be pointlessly complex to explain Inga's actions in terms of beliefs about her memory. ...In an explanation, simplicity is power. (Clark and Chalmers 1998: 13–14)

We do not explain Inga's action in terms of her belief about her own memory, e.g. that she has stored information about the location of the museum in her brain, which she then accesses in order to take action. This would be "pointlessly complex". In normal cases, such as Inga's, we use our own memory transparently:

we do not have to form beliefs about what is stored in our memories. But this is also how Clark and Chalmers describe Otto: he is so accustomed to relying on his notebook that he uses it transparently. He does not have to reflect on where the relevant information is stored; he simply reaches for the notebook. And, all things being equal, we should prefer the explanation that is simplest. Thus, because the extended mind thesis allows us to give a simpler folk-psychological explanation of Otto's actions it is preferable to the embedded account, which requires us take additional explanatory steps (further discussion in Drayson and Clark, forthcoming).

#### 4.2 Usefulness in Theory Selection

One reason to prefer one theory to a rival is its usefulness. As part of a larger discussion on addiction and responsibility, Levy argues that the extended mind thesis is more *useful* than the rival embedded account insofar as it enables us to better control ourselves (Levy 2007a: 220). He argues that research on ego-depletion suggests that addicts have depleted self-control and thus they experience more difficulty in resisting their cravings than one who craves but is not addicted. Levy suggests that “[i]t may be literally impossible for the addict to refrain from taking their drug... when it is immediately available and their self-control resources are depleted.” (Levy 2007a: 219). Nevertheless, he maintains that there are some things that addicts can do both in the short and the long-term to overcome their addiction—namely, they can take steps to control their environment.

According to Levy, the traditional view that the brain wholly constitutes the mind (what we've called intracranialism) works against those suffering from addiction. This view promotes the idea that the only way for one to recover is to change their mind—that is, their brain. In other words, addiction is entirely a matter of “will-power” and the addict needs to just “say no” to their cravings (Levy 2007a: 219–220). Levy argues “[t]o the extent to which we promote the view that giving up a drug, whether it is tobacco or heroin, is all a matter of “will-power,” we direct them away from the kinds of environmental modifications they need to make if they are to regain control.” (Levy 2007a: 220) Thus, Levy reasons that the extended mind view is more useful in so far as it enables us to better control our own behavior, as well as the behavior of others. We should prefer the theory of the mind that yields the most successful strategies with respect to repairing our minds, Levy argues, and thus the “real-world success” of the extended mind thesis is evidence of its truth. He explains, “[k]nowledge is power: if the [extended mind] hypothesis were false, then it would not yield successful strategies.” (ibid.) The stronger constitutive claim that the extended mind makes insists that the just “say no” view about addiction must be rejected.

I think the usefulness of the extended view over the embedded view in this context might be debated. Both accounts point us to the agent's wider environment, beyond the brain, in order to fully *explain* the agent's decisions and reasoning. So perhaps the constitutive claim is not necessary. Thus, this reason on its own may not be sufficient for preferring the extended view to the embedded view—it is also not obvious that Levy intends it to be. But if we accept Levy's point, then the usefulness of the extended account gives one reason (even if not an infeasible one) to prefer it to the competing theory.

### 4.3 Explanatory Power in Theory Selection

I argue that the embedded view suffers from explanatory impotence, while the extended mind thesis has explanatory power and that this gives us reason to prefer the latter to the former. Explanatory power refers to the ability of a theory to effectively explain phenomena that pertain to its subject matter. Consider an example from Drayson and Clark (forthcoming) of a sub-population of inner-city Alzheimer's sufferers in St. Louis who scored dismally on standard tests, such as the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) protocol. Based on their test scores, the patients should have been living in full-care hospitals. Yet, they were able to cope with the demands of daily life and to successfully live alone in the city. The patients puzzled doctors and Alzheimer specialists. After visiting their homes, however, it was revealed that they had transformed their living environments with many personalized cognitive tools, props, and aids: from message centers, open notes about what to do and when, to labels and pictures on the walls, including labelled photos of family and close friends, and 'memory books' that recorded new events, meetings, and plans. Some had open storage spaces that kept crucial items, e.g. kitchen tools or chequebooks, visible, rather than requiring memory of where these things were kept.

The problem is that the standard tests for Alzheimer's disease rest on the assumption of intracranialism: they only evaluate one's internal memory. The extended mind thesis is able to explain how these patients continued to effectively function in the world. While the (likely implicit) assumption of intracranialism left experts confused about how to explain this phenomenon.

What is more, if we took the tests of internal memory as the only standard, these patients would have likely been forcefully removed from their homes and re-located to controlled hospital settings much sooner than might be necessary. Drayson and Clark point out that the re-location of Alzheimer's patients is often a fateful turning point in which their conditions become more severe. On the extended view this is explained by the fact that this kind of re-location is on par with the infliction of new brain damage upon the patients: in one fell swoop it removes them from the cognitive tools, props, and aids which were supporting their intracranial cognitive capacities. The embedded view, on the other hand, struggles to explain why the change in environment so often leads to a dramatic degradation in the condition of Alzheimer's patients.

In this section I have pointed to three preferable characteristics of the extended mind view—simplicity, usefulness, and explanatory power—that together provide us a basis for preferring it to the competing embedded view of the mind. In the next section I consider several ethical reasons for preferring the extended account.

## 5. Ethical Reasons for Preferring the Extended Mind Thesis

In this section I discuss three ethical reasons that have been put forth for preferring the extended account. First, that the extended mind thesis better protects against harm to the mind. Second, that it better accounts for compensatory rehabilitation



as a way of repairing the mind. And third, that it better addresses concerns about the way we assess the capacities of learning disabled individuals.

### 5.1 Protection of the Mind

Adopting the extended mind thesis can help us better protect our minds from harm. Levy (2007a, b), for example, argues that insofar as we view the tools and technologies we use as a part of the mind, we are more likely to protect those tools from being harmed or removed, in just the way we protect the biological agent, especially the brain, from harm. Levy advances the ‘Ethical Parity Principle’ (EPP) to capture this idea. His principle is meant to complement Clark and Chalmers’s (1998) original parity principle. In its strongest form, EPP is stated as follows:

EPP (strong): Since the mind extends into the external environment alterations of external props used for thinking are (*ceteris paribus*) ethically on par with alterations of the brain. (Levy 2007a: 61)

According to this principle, any alteration or *harm* to external cognitive tools should be treated as ethically equivalent to alterations of internal cognitive tools—namely, the brain. Stealing Otto’s notebook would be morally wrong insofar as it would be theft. But, according to EPP we should view this action as much worse than that, ethically speaking, precisely because the notebook plays an important role in Otto’s cognitive life. The notebook stores important information that Otto regularly relies on: it contains his memories, his beliefs, and so on. Thus, the notebook has a cognitive status, and moral status, equivalent to Otto’s brain. For this reason, stealing Otto’s notebook would be more on par with kicking Otto in the head, causing him serious cognitive damage, than it would be to stealing some property, like his gym bag. According to EPP, we should view theft of a cognitive tool much more seriously than theft of property more generally.

Blitz (2010) argues for a *legal* parity principle, which maintains that when external processes are functionally equivalent to internal cognitive ones we should also treat them as legally equivalent. On his view, the law should protect our cognitive tools from harmful alterations in just the way that it has protected our brains against harm. Just as we have helmet laws that mandate protective devices for our brains, for example, we should also have protections against kinds of external cognitive damage.

As before, we might take the “real-world success” of the extended mind thesis to better protect our minds from harm as evidence of its truth. However, although Levy (2007a) endorses the extended mind thesis, he concedes that the defender of the embedded mind is able to offer equal protections of the mind. The embedded mind theorist might reject the strong version of EPP, but she could account for ethical parity with a weaker principle:

EPP (weak): Alteration of external props are (*ceteris paribus*) ethically on par with alterations of the brain, to the precise extent to which our *reasons* for finding alterations of the brain problematic are transferable to alterations of the environment in which it is embedded. (Levy 2007a: 61)

The weaker version of the principle maintains that alterations to external cognitive tools are ethically on par with alterations to the brain only insofar as the reasons we have for objecting to the latter are applicable to the former. According to Levy, weak EPP would allow the defender of the embedded account to offer equal protections to external cognitive tools without conceding the stronger claim that external props are actually partly *constitutive* the mind. Weak EPP still requires that we treat interventions and alterations to internal (brain) and external operations on a par, unless we can find ethically relevant reasons for drawing a distinction between them. Thus, according to Levy, appealing to protecting the mind does not itself establish sufficient normative reasons for preferring the extended mind to the embedded mind, precisely because both frameworks offer equal protections. Simply recognizing the importance of external resources for cognition, as the embedded theory does, is sufficient to establish the same protections (King 2016: 47–50).

## 5.2 Cognitive Rehabilitation: Repairing the Mind

Drayson and Clark try to respond to Levy's claim that the embedded mind would offer the same protections and would have the same ethical implications more generally as the extended mind thesis. They argue that the extended mind thesis has more significant ethical implications than the embedded view, focusing on how these views offer differing accounts of the cognitive rehabilitation of neuroatypical individuals and of our understanding of cognitive impairment.

Cognitive rehabilitation refers to the process of improving an individual's impaired ability to process and use information. Drayson and Clark explain that there are two dominant strategies of cognitive rehabilitation: *restorative* and *compensatory*. *Restorative* strategies aim at restoring the damaged neural area or circuits. *Compensatory* strategies, on the other hand, aim to achieve the same functional results as restorative strategies but in different ways, for example, by using photos or labels to assist memory. This typically involves adaptive strategies that rely on both internal and external resources to improve information processing and use. Some neuroscientists favor restorative strategies and view compensatory strategies only as a recourse, necessary because of our limited understanding of and access to the brain. On this view, as neuroscience advances, compensatory strategies will eventually be replaced in favor of restorative ones. Thus, while compensatory rehabilitation can be an effective substitute for neural restoration, it does not truly restore the mind—only repairing the damaged or affected neural areas can *restore* the mind. This follows from what I will call the *principle of intracranialism*, which holds that neural activity entirely determines mental activity (from King 2016: 55.) Given its commitment to the position of intracranialism, the embedded account is committed to this principle, while the extended account rejects it. And it follows from the principle of intracranialism that only by restoring neural activity can we restore mental activity.

As a result, Drayson and Clark argue that the extended mind thesis and the embedded mind thesis take different positions on rehabilitative strategies. Due to her intracranialist commitment, the embedded mind theorist must maintain that compensatory strategies cannot achieve true restoration of mental function. So she must maintain a distinction between restorative and compensatory rehabilitation: if

the mind is realized only by neurons, then the only way to repair a damaged mind is to repair the neural areas that bring it about. The extended mind thesis, on the other hand, offers a different picture. It allows us to view compensatory strategies as on par with restorative ones. For the extended mind theorist, both strategies are legitimate ways of repairing cognitive ability—neither is a second-best option.

The oddity of what the embedded mind view is committed to in maintaining the traditional distinction between restoration and compensation is brought out by an example that Clark gives (in response to Jerry Fodor's (2009) critique of the extended mind thesis):

[I]magine a case in which a person (call her Diva) suffers minor brain damage and loses the ability to perform a simple task of arithmetic division using only her neural resources. An external silicon circuit is added that restores the previous functionality. Diva can now divide just as before, only some small part of the work is distributed across the brain and the silicon circuit: a genuinely mental process (division) is supported by a hybrid biotechnological system... (Clark 2009)

In this case Diva's damaged neural circuit has been restored with silicon-based functional replacements of neurons. But, if one maintains the traditional distinction between restoration and compensation, not even this would count as true restoration. If neurons alone can constitute cognition, then even this rehabilitative strategy falls short of true restoration. Even though Diva's doctors have addressed the structural integrity of her neural circuits they have had to rely on silicon structures, rather than biological structures, to do so and thus this has to be seen as merely compensatory. So, while the extended mind theorist can describe Diva's rehabilitative strategy as truly *restoring* her cognitive functioning, the embedded mind theorist cannot.

As another example, consider again the Alzheimer's patients discussed before. These patients developed their own compensatory strategies that allowed them to live successful lives despite their impaired intracranial functioning. The extended mind thesis allows us to see the compensatory strategies developed by these patients as genuinely restoring their cognitive functioning. It offers a new way of conceptualizing the distinction between restoration and compensation, suggesting that we should evaluate rehabilitation based on the functional capacities of the extended cognitive system, rather than restricting cases of "true" rehabilitation to only those that involve restored neural circuits.

Drayson and Clark's argument also offers a response to Levy's view that the embedded mind thesis has the same ethical implications and affords the same protections of the mind as the extended mind thesis. Drayson and Clark argue that if we view the cognitive tools and aids of the cognitively impaired as mere scaffolding, rather than as legitimately a part of their minds, we are *less likely* to protect those tools from harm in the same way that we protect the biological brain from harm. On the extended mind view, the minds of people who rely on external tools are more vulnerable to harm than those who do not rely as heavily on them precisely because the dominant intracranialist view of the mind has failed to recognize and protect those tools. Thus, resisting the extended mind thesis may lead us to neglect what ought to be protected and, as a result, place cognitively impaired people in greater danger of cognitive harm (further discussion in King 2016: 46–48).

### 5.3 Capacities of Learning Disabled Individuals: Improving the Mind

Finally, King (2016) offers yet another reason to favor the extended account: that adopting the embedded account commits one to problematic views about the cognitive capabilities of learning disabled individuals, while the extended account avoids these commitments. Researchers working on learning disabilities draw a distinction between two kinds of strategies aimed at addressing learning that, according to King, roughly maps on to the distinction between restorative and compensatory strategies discussed in the previous section. Researchers of learning disabilities distinguish between “remedial” strategies and compensatory strategies (e.g., Garner and Campbell 1987). Remedial strategies aim to directly address a learning-disabled individual’s impairment by improving their ability to perform tasks in just the same way that a non-disabled individual would (King 2016: 54). Compensatory strategies, on the other hand, attempt to circumvent learning impairment by helping the individual perform the same tasks by using assistive technologies (ibid). King argues that remedial strategies are analogous to restorative strategies in cognitive rehabilitation and (again, because of the commitment to intracranialism) the embedded mind theorist has to say that remedial strategies are the only “true” way to enhance a learning-disabled person’s cognitive capabilities. Compensatory strategies, as before, are only a second-best option, employed when remedial strategies are not possible: they might help an individual *compensate* for her impairment but they do not *restore* cognitive capacities. The extended mind theorist, on the other hand, can view both strategies as genuinely restoring cognitive capabilities.

To bring out this difference, King describes a “paradigm case” of a learning disabled person who uses assistive technologies as a compensatory strategy for her disability:

Consider someone with a learning disability, Dana, who requires a graphic organizer of potential decisions in order to evaluate which decision is best. In this example, let us imagine that Dana has a very difficult time comparing the relevant factors when she must evaluate them solely “in her head,” but when allowed to create and utilize a visual diagram of the various possibilities, her decision making skills are just as good as anyone’s. In this case, she needs a particular physical configuration of information in order to be able to perform a cognitive process like comparing and choosing among potential courses of action, and without the aid of these external resources, it would appear that she is incapable of performing this cognitive action. (King 2016: 49)

According to King, environmental tools, such as graphic organizers (i.e. visuo-spatial ways of representing information such as cognitive maps, venn diagrams, flowcharts) are typically employed as assistive technologies in compensatory strategies for learning, problem solving, and planning, for example. The embedded mind thesis would say that Dana couldn’t make complex decisions. King argues that we should resist this conclusion. Dana may need to rely on graphic organizers, but she is quite capable of making complex decisions. She just requires a particular sort of environmental scaffolding that a non-learning disabled individual would not. In other words, Dana is only unable to make complex decisions when she is denied access to the external cognitive aids that she requires (King 2016: 50). Thus, the extended account better accommodates our intuitions about learning-disabled individuals than the embedded account.

### 5.3.1 *Response from the Embedded Mind Theorist*

Is this an accurate portrayal of the embedded mind theory? One might object that the embedded mind thesis can allow that compensatory strategies genuinely repair and improve the mind. After all, the embedded mind theorist holds that the environment plays a crucial and sometimes indispensable role in supporting cognition. Thus, just because the embedded mind theorist denies the constitutive claim that the extended mind theorist endorses does not mean that she would also resist compensatory strategies that make use of assistive technologies (King 2016: 54–55). But given the commitment to the principle of intracranialism, the embedded mind theorist is committed to an inverse relation between the extent to which an individual relies on external tools and the extent to which we ought to say that she, or her mind, is really doing *x*, where *x* is some cognitive process (ibid.). This means that Dana is only “doing” as much, cognitively speaking, as her neurons are doing (King 2016: 56). And therefore, she only merits “cognitive credit” for what her neurons do: the cognitive work that is done by whatever assistive devices she employs is not being done by her and so she should not get credit for the achievements of these devices. King points out that this means that the more heavily integrated assistive devices are—the more one relies on these tools—the less cognizing one is actually doing. The embedded mind theorist is, thus, forced to say that Dana has less cognitive capacities and deserves less cognitive credit than someone who could perform the same task ‘intracranially’. So, the embedded mind thesis really does commit one to saying that learning-disabled individuals, such as Dana, who rely on assistive technologies are cognitively capable of less than non-disabled learners.

On the extended view we need not draw equivalence between neural capacity and cognitive capacity, and, thus, diminished neural capacity does not entail diminished cognitive capacity. This means that compensatory strategies are not just acceptable (as the embedded mind theorist would also accept them), but they are equally legitimate ways of improving and repairing cognitive capacities as restorative strategies are (ibid.). King argues that the extended view better captures our intuitions about learning disabled individuals: “[a]ssistive technologies are tools that help [learning-disabled] individuals do more, not less” (King 2016: 57). But furthermore, citing various testimonial reports, she also argues that the extended view better captures how learning-disabled individuals view themselves and their own relationship to the assistive technologies that they use:

[Learning-disabled] individuals who have consistent access to assistive technologies in the classroom report feeling less anxious, more independent, and more confident in their own abilities (Day and Edwards 1996). Rather than making them feel as if the more they use technologies, the less they can do “themselves” (the picture of cognitive capability that the embedded thesis predicts), well-assisted [learning-disabled] individuals report feeling as if they are capable of doing more, and with an increased sense of independence and self-reliance. Testimonial reports of the impact of access to assistive technology on [learning-disabled] individuals’ self-concept further supports the suggestion that [learning-disabled] individuals see themselves in this way. (King 2016: 57).

While this is hardly an indefeasible reason to favor the extended view, the first-person accounts from learning-disabled individuals do seem to align with our

own intuitions that assistive technologies help them to become better at processing information and using information.

The embedded mind defender might also maintain that Dana's graphic organizer helps her arrive at better decisions than had she not used the device. But, there is a key difference between what, precisely, the two accounts mean when they say that the technology allows Dana to "do more": the embedded view would say that the graphic organizer enables Dana to achieve more qua competent tool user, while the extended view asserts that the graphic organizer allows Dana to do more qua cognizer (ibid). King uses the analogy of a runner who would fatigue after just a few miles when running barefoot but when given a pair of quality running shoes (an assistive device) could run for significantly longer distances and at greater speeds. The shoes could be seen as merely an assistive technology—a "tool" for running—but what we want to say, and what we are more likely to say, is that their use improved the runner's abilities qua runner (King 2016: 60). Likewise, the extended thesis views the cognitive agent and the tool as a single (wide or extended) system: it draws no divide between the agent and the tool, while the embedded thesis would have to attribute any increased capabilities that result from the tool-use to the tool itself rather than the individual. Because the embedded theorist clings to intracranialism, she must insist on a clear distinction between the cognitive capacities of the agent and the non-cognitive capacities of the tool (King 2016: 59–60). King's argument suggests that while the extended mind thesis might seem radical and less intuitive, on a closer analysis of how the view explains rehabilitation strategies employed by learning-disabled individuals it seems it is the embedded thesis that struggles to accommodate our intuitions about the use of assistive technologies. Furthermore, the embedded view must reject the beliefs that learning-disabled individuals self-report about their own relationships to their assistive devices.

I conclude that King's argument offers another normative reason to prefer the extended account. The ethical reasons discussed in this section should also serve as caution that our biological capacities should not condition our notions of rehabilitation, well-being, or decision-making, at the risk of alienating, marginalizing, and even harming those who rely on technologies for their cognition. The picture that I hope emerges from this paper is that technology can, and often does, allow our minds to transcend our biological capacities.

## 6. Conclusion

The debate between the embedded and the extended mind views has seemed to reach a stalemate. Is the mind merely embedded in the world, coupled to the various tools that it uses, or is it partly constituted by these tools? And what hinges on this? Insofar as one agrees that we have reached a stalemate, I have argued that there are normative reasons that we can appeal to that make that extended mind thesis preferable to the embedded mind thesis. First, in terms of choosing between these views as two scientific theories of the mind, we can appeal to various norms of theory selection. I have argued that the extended mind thesis has three virtues—simplicity, usefulness, and explanatory power—that make it preferable to the embedded mind view. Second, looking at the ethical import of these views, I argue that

there are again good reasons to prefer the extended mind thesis: it better protects against harm to the mind, it better accounts for how we repair the mind, and finally, it offers a better assessment of the capacities of learning disabled individuals. While some of these are not on their own decisive reasons to prefer one view to the other, I believe that taken together, they provide good reason for preferring the extended view. Thus, taken as a whole these reasons can be used as a response to the coupling-constitution deadlock. We should, I argue, view external tools as enabling us to transcend our biological capacities: our minds are extended, not just embedded.

#### References:

- Adams, F. and K. Aizawa. (2001), "The Bounds of Cognition." *Philosophical Psychology* 14(1): 43–64.
- Adams, F. and K. Aizawa. (2008), *The Bounds of Cognition*. Blackwell Publishing.
- Blitz, M. J. (2010), "Freedom of Thoughts for the Extended Mind: Cognitive Enhancement and the Constitution", *Wisconsin Law Review* 4: 1049–1117.
- Bratman, M. E. (1987), *Intentions, Plans, and Practical Reasons*. Center for the Study of Language and Information Publications.
- Clark, A., and D. Chalmers. (1998), "The Extended Mind." *Analysis* 58(1): 7–19.
- Clark, A. (2008), *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. New York: Oxford University Press.
- Clark, A. (2009), "Response to Jerry Fodor's Review 'Where Is My Mind?'" Review of *Supersizing the Mind*. *London Review of Books*: Letters 31(6).
- Clark, A. (2010). "Coupling, constitution, and the cognitive kind: A reply to Adams and Aizawa". In: R. Menary (ed.), *The Extended Mind*. Cambridge: MIT Press, pp. 81–100.
- Drayson, Z., and A. Clark. (Forthcoming). "Augmentation, Agency, and the Spreading of the Mental State." <<https://sites.google.com/site/zdrayson/papers>>
- Drayson, Z. (2010). "Extended Cognition and the Metaphysics of Mind", *Cognitive Systems Research* 11(4): 367–377.
- Garner, J. B., and P. H. Campbell. (1987), "Technology for Persons with Severe Disabilities: Practical and Ethical Considerations", *The Journal of Special Education* 21(3): 122–132.
- Levy, N. (2007a), "Chapter Six: The 'Self' of Self-Control", *Neuroethics: Challenges for the 21<sup>st</sup> Century*. Cambridge: Cambridge University Press, pp. 197–221.
- Levy, N. (2007b), "Rethinking Neuroethics in the Light of the Extended Mind Thesis." *American Journal of Bioethics* 7(9): 3–11.
- Levy, N. (2011), "Neuroethics and the Extended Mind". In: J. Illes et al (eds.), *Oxford Handbook of Neuroethics*, Oxford: Oxford University Press, pp. 285–294.
- King, C. (2016), "Learning Disability and the Extended Mind" *Essays in Philosophy* 17(2): 38–68.
- Kuhn, T. (1977), *The Essential Tension. Selected Studies in Scientific Tradition and Change*, Chicago: University of Chicago Press.
- Rowlands, M. (2010), *The New Science of the Mind*. Cambridge, MA: MIT Press.
- Rupert, R. D. (2004), "Challenges to the Hypothesis of Extended Cognition", *Journal of Philosophy* 101(8): 389–428.
- Shapiro, L. A. (2008), "Functionalism and the Boundaries of the Mind", *Cognitive Systems Research* 9(5): 5–14.
- Simon, H. (1969), *The Sciences of the Artificial*, Cambridge, MA: MIT Press.
- Sprevak, M. (2009), "Extended Cognition and Functionalism", *Journal of Philosophy* 106: 503–527.
- Vygotsky, L. (1930), *Mind in Society: The Development of Higher Psychological Processes*. M. Cole, V. John-Steiner, S. Scribner, and E. Souberman (eds), Cambridge, MA: Harvard University Press.

## Karina Vold

### Prevazilaženje zastoja: naučni i etički razlozi za prihvatanje teze o proširenom duhu

#### Apstrakt

Prema tezi o proširenom duhu iako se naši umovi prevashodno nalaze u našem mozgu i telu, njih ponekad delimično ustanovljuju delovi naše sredine. Kritičari tvrde da nemamo razloga da od teze da je naša kognicija uronjena u našu sredinu napravimo korak ka prihvatanju jače teze prema kojoj je naša kognicija konstituisana našom sredinom. U ovom radu, pokazujem da postoje normativni razlozi, naučni i etički, da prihvatimo tezu o proširenom duhu umesto suparničke, uronjene, teze.

**Ključne reči:** proširena kognicija, intrakranializam, uzglobljen duh, etika, selekcija teorija, kognitivna rehabilitacija