Tools, Language and the Aggregate Mind

Paper for the course Philosophy of Artificial Intelligence 2008
Matthijs Westera, studentnr 3021211

Abstract. This paper addresses an important open question in cognitive science: what is it that makes us human? Two uniquely human capacities will be discussed. The first is tool-use, which we will discuss initially from the perspective of Clark and Chalmers' *Extended Mind Hypothesis* (EMH). Some problematic aspects of the EMH are highlighted and a view on tool-use based on Dennett's *homuncular functionalism* is introduced as an alternative. The second capacity is language, which we will discuss from a memetic viewpoint. We formulate the *Aggregate Mind Hypothesis*, stating that the human mind is not identical to the software implemented on the brain, but rather to an aggregate mind of the brain and the vehicles of language.

> "The ultimate irony of human existence is that we are supreme individualists, whose individualism depends almost entirely on culture for its realization." (Merlin Donald (2001), *A Mind So Rare*, p12).

1 Introduction

What is it that makes us human? On the one hand human beings are quite unlike any other species on earth, but on the other hand neuroscientists mostly find correspondences, rather than differences, between the human brain and that of monkeys. Although the basic areas responsible for perception and motor control have been found, we cannot seem to localize the human capacities such as consciousness, declarative memory and beliefs. How can such a primitive brain generate the richness of the human mind? In this paper we will look at two uniquely human capacities in order to shed light on the issue.

The use of tools is a capacity that is nearly uniquely human. No other animal species governs such a rich arsenal of especially manufactured, complex tools. Tools, in turn, have changed in many ways the way we live and the way we think. To such an extent, even, that Clark and Chalmers (1998) among others claim that cognition no longer resides only in our heads, but extends into the world through tools. This is the Extended Mind Hypothesis (EMH), which we will summarize in section 2. In section 3 we will raise some objections to the EMH and propose an alternative view more explicitly centered on tool-use, from the viewpoint of Dennett's homuncular functionalism (1987).

Another uniquely human capacity is language. In section 4 we discuss the role of language in cognition and formulate the Aggregate Mind Hypothesis (AMH), stating that the human mind is an aggregate of brain and language. Section 5 contains some concluding remarks.

2 The Extended Mind Hypothesis

Consider Inga and Otto, both wanting to visit the Museum of Modern Art (this is the original example from Clark and Chalmers 1998). Inga correctly believes the museum is on 53rd Street, goes there and enters the museum. Otto, on the other hand, has Alzheimer's disease and cannot rely on biological memory. Instead, Otto always carries a notebook with him, in which he writes down any new information he encounters. Otto furthermore always consults the notebook before taking action.

Clark and Chalmers claim that when Otto's notebook says the museum is on 53rd Street, Otto clearly *believes* the museum is on 53rd Street just like Inga. Since Inga's biological memory and Otto's notebook fulfill the same functional roles, they are the same kind of mental process, even though one is 'extended' in the sense that it is implemented outside the brain.

As illustrated by this classical example, the EMH claims that cognitive processes and ultimately the mind are not bound by flesh and bones. Some functions that might usually be implemented on neurons, can be implemented

on external objects as well, without them being excluded from the realm of the cognitive. A term used to describe this is *active externalism* (as opposed to *passive* or *meaning externalism*), indicating that the environment plays an active role in cognitive processes.

The EMH has been formulated in different ways. For example, Rowlands (2003) describes the EMH as *vehicle externalism*, expressing that the cognitive enabling system or vehicle for cognition may be implementationally extended. Clark and Chalmers themselves formulated part of the EMH in the *Parity Principle*:

> **Parity Principle** (from Clark and Chalmers (1998))
> 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.

In their formulation of the Parity Principle, Clark and Chalmers included a parenthesized "for that time" indicating a notion of dynamicity and temporariness in the implementation of cognitive processes. When talking about an extended *mind* rather than extended cognition, they counter this dynamicity via a number of criteria, and instead hold on to the mind as a process implemented on a static vehicle.

Sticking largely to the formulations of Clark (to appear), these criteria are (i) the *automaticity criterion* that the resource must be reliably available and typically invoked, (ii) the *reliability criterion* that any information obtained from the resource must be more-or-less automatically endorsed and (iii) the *accessibility criterion* that information contained in the resource should be easily accessible as and when required. Otto's notebook can be seen to meet these criteria, whereas, for example, other people's opinions usually do not meet any of the criteria.

The EMH presupposes a functionalist approach to the mind, claiming foremost that what does and does not count as a cognitive process does not depend on the implementation. Functionalism, to put it boldly, claims that mental states are determined by the functional roles they fulfill, so they may as well be implemented on a silicon chip instead of neural tissue, and equally well on a combination of neural tissue and extensions such as a microchip implant or, more old-fashionedly, a simple notepad. So, in light of the traditional functionalist approach, the Parity Principle does not seem to make any too strong claims.

3 Homuncular functionalism and tool-use

From this point onwards we will take the specific viewpoint of *homuncular functionalism* as advocated by Dennett (1987), which we will summarize (and slightly expand for our purposes) first. Then we will highlight three disadvantages of the current form of the EMH. At the end of this section, an alternative view is proposed.

Homuncular functionalism is a branch of functionalism that attempts to disassemble the mind into parts or modules that are functionally simpler, and to do this recursively down to the level of very elementary modules. These elementary modules are then so simple that is no longer difficult to see how they are implemented, thus solving the problem of how the brain realizes the mind. In Dennett's view, the mind can be disassembled into modules that are themselves full minds, with mental properties and a certain degree of intelligence. Dennett postulates the existence of a hierarchy of minds that become more stupid as we go down in the hierarchy. Eventually the simplest minds would be so simple that their vehicles, for example their implementation in neurons, silicon chips or towers of matchboxes, can be easily understood. An exemplary decomposition of the human mind is depicted in Figure 1.

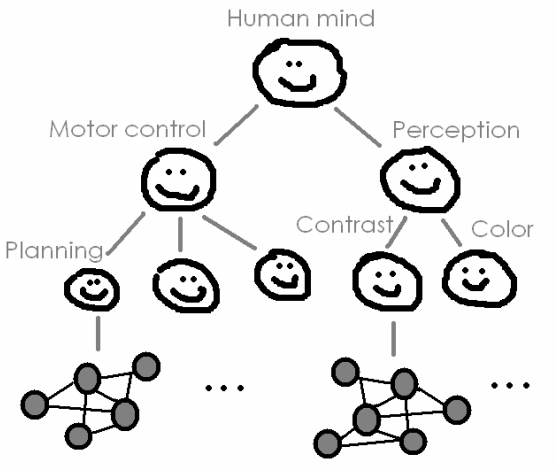

Figure 1. A tentative homuncular decomposition of the human mind.

We have chosen in this paper to view the EMH and, later on, its alternative, from the perspective of homuncular functionalism because it is one of the branches of functionalism that makes a concrete attempt at explaining how the brain realizes the mind, by linking the lowest level of homunculi directly to implementation. This does not mean that in the eyes of the homuncular functionalist the implementation of a cognitive process is limited to, for instance, neurons. It is still a functionalist view, for the same functional module may be implemented on various different substances. But since implementation is not neglected, homuncular functionalism is fit for the job. As we will see, this perspective leads to a conceptually elegant way of dealing with the use of tools.

Let us have a look at the EMH from the viewpoint of homuncular functionalism. According to the EMH, existing modules of the mind are replaced by functional equivalents that are implemented elsewhere, possibly outside the skull. This extended mind is still a human mind. The process is depicted in Figure 2, in the hierarchical fashion of homuncular functionalism.

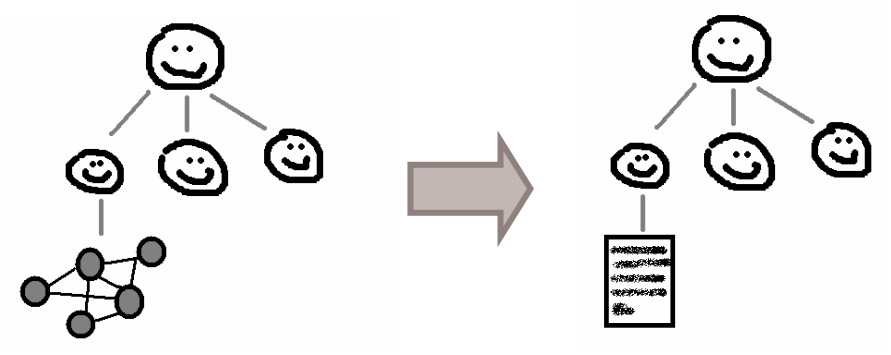

Figure 2. The replacement of a homunculus or functional module by a functional equivalent which may be implemented externally.

In our opinion, the EMH in this form is too liberal. A true replacement of a module in the mind can occur only when some object has the same causal connections with other modules in the mind as the module it is meant to replace. Tool-use, however, is always mediated by sensorimotor areas in the brain. Since perception and motor control are functional modules themselves, a tool mediated by these modules cannot have the exact same set of causal connections as a module which is not.

The famous example of Otto and his notebook seems to appeal to a behaviorist rather than a functionalist intuition, namely that reading in a notepad that the museum is on 53rd Street causes the same behavior as retrieving from biological memory that the museum is on 53rd Street (Otto walks towards that particular street), *ergo* both mental processes are of the same type. However, for Otto the process that we mistakenly call "remembering" is, quite unlike true remembering, always mediated by a mental state of having read that the museum is on 53rd Street. This sensorimotor mediation of supposed mind extensions is a problem for the EMH.

Another objection to the EMH is its focus on compensating cognitive deficits, as in the case of Otto's memory. Although the Parity Principle does not rule out the possibility of extensions of the mind with processes that never occur in our heads, it does not explicitly leave open this possibility either. Like for instance Wheeler (draft) we feel that a healthy mind is perfectly able to

implement a cognitive process together with an external instrument, even when the instrument's added functionality could never be implemented in the head.

Following the first objection, Otto and his notebook can be considered a specific instance of the more general phenomenon of a human mind using a tool. There is no actual replacement going on as depicted in Figure 2. Rather, Otto's original brain-implemented mind simply lacks a memory module and forms an aggregate, *higher level mind* together with a notepad to be able to achieve roughly the same. We will elaborate on this aggregation process soon. One could say this objection applies to cases outside the intended scope of the EMH, but since tool-use is such an everyday business and since it seems a generalization of cases like Otto and his notepad, the objection is relevant nevertheless.

There is a third objection to the EMH. Let us reconsider the status of the mind as regulated by the EMH and the automaticity, accessibility and reliability criteria. This has been discussed in detail by for instance Rupert (2004). The criteria are meant to safeguard the status of the mind as something constant, something which is not dynamically extending and shrinking as one interacts with different tools. However, as Rupert notes, in modern society mobile telephones with a system of directory service are always reliably available to most people, typically invoked when required, and the phone numbers retrieved from the service are automatically endorsed. Does this mean that Otto, who happens to carry a mobile phone along with his notepad but who has never called you, knows your telephone number – or that Inga does so, for that matter? Similar objections hold with respect to the omnipresence of internet access to most university students and the automatic endorsement of anything Wikipedia says.

The three criteria leave unwanted room for absurd kinds of mind extensions. To solve this, Clark and Chalmers have formulated a fourth criterion, the *past endorsement criterion*, that any information retrieved from the instance must have been consciously endorsed in the past. But this fourth criterion leads to the acceptance of an internal privilege – i.e. there is something special about

internal cognitive processes – contrary to the functionalist view (and especially the EMH). In addition, it wrongly denies the existence of beliefs (also in a biological brain) that have never been consciously endorsed (Rupert 2004).

We have given three main objections to the EMH. In the remainder of this section we will argue for an alternative, grounded in homuncular functionalism and avoiding the problems mentioned.

Any human being might use a tool to do things the person alone could never do in his head. For example, to the human mind a handheld computer adds functionalities such as a high-precision calculator, an unlimited long-term planning capacity, the possibility to download and upload an abundance of information and a photographic memory. Instead of considering homunculus replacement as in the original EMH, we claim that the entity or mind responsible for any complex process involving both the human mind and the handheld computer (or any other tool) is an aggregate of the person's mind and functional modules implemented on the instrument. An aggregate mind in this sense is depicted in Figure 3, consisting of a brain-implemented mind (left) and the functionality of some other device.

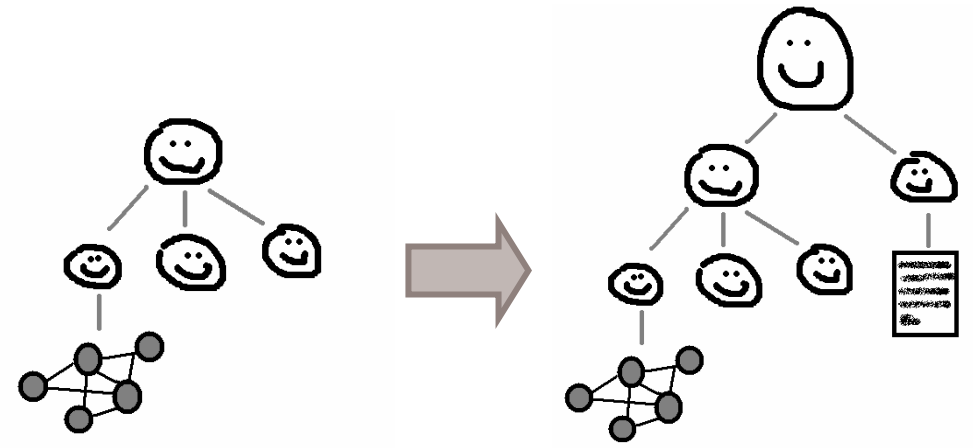

Figure 3. The aggregation of minds or functional modules, into bigger minds or modules. Here one module is implemented on a notepad.

According to homuncular functionalism (and contrary to the EMH with its criteria) the vehicles of cognition are dynamic entities, as different aggregate minds constantly appear and disappear when a human mind interacts with different instruments, computers and other minds. It seems that this dynamicity does not comply with the strong intuition that one mind in particular belongs to a person. This is where the criteria come in, but only as pragmatical rules of thumb. If a person is always seen with a certain external instrument (accessibility), always consults it before taking action (automaticity) and trusts it (reliability), then chances are that we start talking about the higher-level mind, that is implemented on the person plus the instrument, as the person's mind. In other words, when the criteria are fulfilled, we move up in the homuncular hierarchy.

In our opinion, if we reduce the criteria to mere pragmatic ones rather than ontological ones, their vagueness and exceptions as summarized above are no longer as problematic. Pragmatics is traditionally more used to dealing with vagueness after all, and most exceptions (such as a person knowing *every* telephone number) can be ruled out on the basis of absurdity. Interestingly, we can see how the criteria, as pragmatical rules of thumb, derive naturally from one of pragmatics' favorite principles, that of *cognitive economy*. It is economical to stay at the highest reliable level in the homuncular hierarchy, because the higher the level, the lower the number of homunculi postulated to explain the world. Of course, constantly adjusting to dynamically appearing and disappearing higher-level minds is ineffective. So only when a higher-level mind becomes a stable entity, which is approximately the case when the criteria are fulfilled (give or take some absurd exceptions), we can shift to this higher, more economical level. This explains why we view Otto's mind as encompassing both Otto's brain and his notepad, and why we feel Otto is missing something when he is seen without the notepad.

In short, mind-aggregation is what happens when a brain-implemented mind uses a tool, and we shift to this higher level when the tool-use complies with the

criteria. Admittedly, the view presented here somewhat stretches our notion of a mind, but this stretching has already been done by Dennett. We merely apply his ideas to the direction of growing complexity, upwards in the hierarchy. As we explained before, Otto and his notebook can be seen as an instance of homunculus aggregation as in Figure 3 (with the left homunculus simply lacking a memory module), rather than homunculus replacement as expressed in the EMH. To give another famous example (but from a slightly different context), the *systems reply* to the Chinese Room argument (Searle 1980), which roughly claims that the room as a whole understands Chinese, appeals exactly to this notion of an aggregate mind.

4 Language and the human mind

The previous sections have paved the way towards accepting an aggregate, extended mind which is, guided by the (pragmatical) criteria, still considered a human mind. In this section we will argue that one typical functional extension of the brain-homunculus is language, together with the brain-homunculus forming the conscious mind that we think of as human. Because we wish to stick to homuncular functionalism, we should somehow view language as a homunculus or functional module on its own. Language is a very complex, dynamic and distributed phenomenon, so it is not immediately trivial how to view it as a functional module. In this section we will first show that viewing language as a *memeplex* enables us to do just that. After that we will explore the link between language and the human mind.

Memes, as coined by Dawkins (1979), are the cultural equivalents of genes. Although opinions differ, memes are generally considered elementary cultural phenomena, such as trends, habits, words and proverbs. Analogously to biological genes, these memes can be seen to strive for *selfish reproduction*. Fit memes will be reused by human brains a lot and eventually become part of a human culture and less fit memes will become extinct. Susan Blackmore adopts this terminology and organizes it in her 1999 book, introducing the term

‘memeplex’ to denote a set of memes that reinforce each other and replicate together, for example a religion or a culture. Western culture is such a memeplex, as well as the cyber-gothic scene or the Dutch language.

In terms of homuncular functionalism, the memeplex we know as language is a homunculus, a functional module. However, due to the sheer complexity of language as a whole, and because it is a massively distributed phenomenon (not quite unlike, we like to mention nonchalantly, the human brain), a simple input-output relation for this module cannot be found and it seems that attributing mental properties to the homunculus would make things a lot easier. The language homunculus may be described as having a fear of dying out, an intention to reproduce, selfishness, maybe even a trace of memory (in the shape of proverbs, for example, or written text).

When speaking of language in these terms instead of describing the complex memetic dynamics of a language system, it is easier to see how such a system would behave. Of course the solely pragmatic necessity to assign mental properties to a homunculus does not lead to any ontological claims on mental properties. Nor is this pragmatic switch in jargon a requirement for language to be compatible with homuncular functionalism (for it would be absurd to assign mental properties to a notepad or a screwdriver on its own). But the idea itself, that language has these mental properties, should not seem too outrageous anymore. We will soon explain why this is relevant.

If we view language as a functional unit, with or without mental properties, it should be possible to decompose it into smaller units that are eventually grounded in matter. Although according to the functionalist approach any functional module is multiply realizable, it is still interesting to consider in what kind of stuff the language system is implemented. Although, analyzing this in more detail is beyond the scope of this paper, we feel it is safe to state that language is implemented on a massively distributed system, consisting at least of all the human brains and other media in which its words and expressions are somehow represented. This would mean that each human brain fulfills two roles: one as the implementation of the basic sensorimotor capacities of a

human being, and one as part of the entirety of human brains that together with other media implement language. This may seem contradictory, but there is no a-priori reason why the brain cannot fulfill two such roles at the same time. Consider, for example, the hardware of a desktop computer, which functions both as a basic calculator and word-processor and simultaneously as a part of the World Wide Web, two essential ingredients of the great tool we usually denote when we talk about our 'computer'. In this sense the computer is a higher-level aggregate of the software implemented on a desktop and the largely invisible World Wide Web, which is implemented in part on the same desktop.

The capacity to comprehend and produce natural languages with enough expressivity to write a philosophical report is uniquely human. There is still much discussion going on about what enabled early humans, unlike any of the other intelligent mammals on Earth, to develop this capacity. But whatever triggered it, language in turn became an instrument that influenced the nature of our species, from the way we live to the way we think. Extensive research on the role of language in human cognition has been done for example by Tomasello (1999) and Donald (2001).

Language can be seen as a tool used by human brains. As such, it is obvious that from a very early age language is always accessible, used automatically and extremely reliable. It is therefore more economical to describe the world not in terms of the software running on our brains, but rather in terms of the software running on higher-level aggregates of a brain plus the vehicles of language. The human mind, we argue, is the aggregate of the 'animal mind' that allows us to perceive and move, which is implemented on the brain, and the complex memeplex of language, implemented on a distributed web of brains and other media, responsible for higher-level cognitive processes such as declarative memory, symbolic thought and consciousness. In accordance with the cognitive economy principle and the derived pragmatic criteria, this aggregate is what we consider a human mind, on which we project our folk-psychological goals, feelings and beliefs. We will call this the *Aggregate Mind Hypothesis* (AMH).

Aggregate Mind Hypothesis

The human mind is a widely distributed phenomenon. It is the aggregate of a reactive animal mind (implemented on the brain) and a linguistic memeplex (implemented on a complex web of many brains and other media).

The AMH as such is depicted in Figure 4.

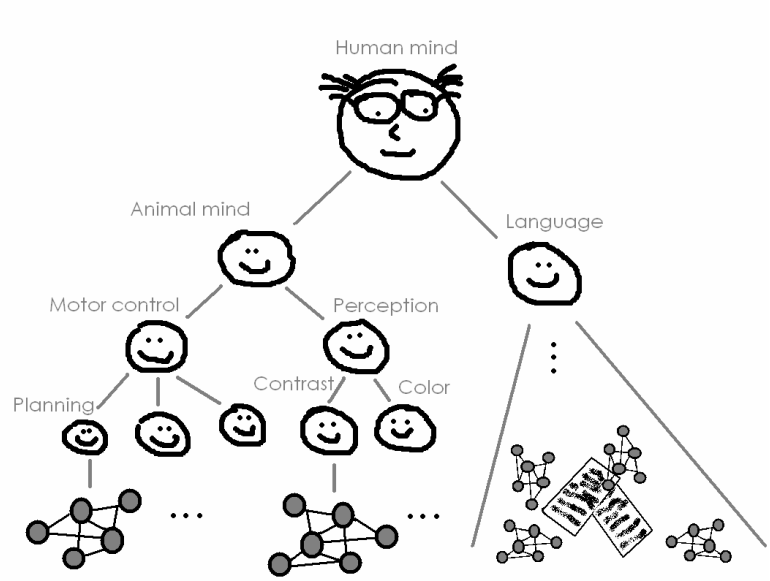

Figure 4. The human mind as an aggregate mind of brain and culture.

Earlier in this section we explained why we feel that mental properties can be ascribed to a language system, if only for pragmatic reasons. Homuncular functionalism attempts to partition the mind in homunculi of decreasing complexity, each of which can have mental properties to a certain extent. By accepting that the human mind is an aggregate of brain and culture, and pursuing the aims of homuncular functionalism, it is only natural to accept that language, as a lower-level homunculus of the human mind, has mental properties just like the supervening human mind itself. The mental properties propagate downwards to the level of neurons on one side and to the distributed mass of brains, paperwork and other media on the other.

5 Conclusions and perspective

We have formulated several objections to the EMH as proposed by Clark and Chalmers (1998). First, we claimed that the EMH is too liberal in assigning the same functional role to external instruments mediated by the auditory, visual or tactile input-output modules, and modules in the brain which require no such mediation. Second, we objected that the EMH deals mainly (though not exclusively) with cognitive impairments that are compensated for by external devices, such as Otto's memory. We feel that a healthy mind is just as capable of combining with extended hardware to implement a cognitive process. Third, we argued that the idea of the mind as a stable entity rather than a dynamically changing one, as regulated by the three criteria, is a pragmatical issue rather than an ontological one, by linking it to the cognitive economy principle. It is more an issue of what we regard as a person than of what minds really are.

We then employed Dennett's homuncular functionalism as an alternative view dealing with the use of instruments in a wider sense than the EMH does. Homuncular functionalism is very liberal with respect to the mind, allowing (but not requiring) basically any functional module or homunculus to be ascribed mental properties when convenient. It relies on pragmatic criteria to establish what we generally consider as the mind of an individual, complying with the intuition of minds as stable entities without making a problem of exceptions like the phone number service or the library.

After introducing the concept of mind aggregation, we showed how language could be regarded as a functional module like any instrument or mind, and encouraged the idea of language having mental properties by describing it as a memeplex. This served the goal of showing that the human mind, as formulated in the AMH, is an aggregate mind of language, implemented in a mass of brains and media, and the software implemented on the brain.

Accepting the AMH is not without consequences. The AMH implies that the search for brain areas corresponding to certain cognitive processes, as

conducted by neuroscientists, may very well be futile in some cases. The reactive capacities we share with other mammals, such as basic perception and motor control, can be (and more or less have been) localized in the brain. But the higher-order cognitive capacities that make us human (e.g. declarative long-term memory, language capacity, the self and consciousness) are typically grounded in the aggregate of brain and culture and can therefore not be localized in the brain alone. Some of the brain areas responsible for the mediation of such processes have been discovered (e.g. the hippocampus for storage and retrieval in declarative memory), but the actual memory traces remain a mystery for the neuroscientist. Only when the human mind is rightly considered a mixture of cultural and biological processes will cognitive neuroscience advance.

Interestingly, Merlin Donald arrived at a similar conclusion from a more neurocognitive and less philosophical viewpoint, which he describes in his book *A Mind so Rare* (2001). He describes for example how abandoned children, who grow up entirely deprived of language, have been found to entertain a much weaker sense of self and a much weaker conscious awareness, comparable to the levels found in non-human primates.

Donald argues that the brain, despite being a piece of extraordinarily complex matter, is only capable of generating a reactive, percept-processing mind like that of a non-human animal. However, when this complex matter is conjoined to an invisible cultural web, the symbolizing powers of the human mind emerge and the range of its own conscious awareness is expanded. The human mind, he argues, is a *hybrid mind*. He concludes his book with the following remark on the strangeness of this notion:

> "If this appears strange to us, this is surely only a reflection of our conventional notions of, among other things, strangeness. We have lived comfortably with the myth of the isolated mind throughout most of our history. [...] The triumph of consciousness will be

complete when it can finally reflect on the collective process itself and see only itself, in the mirror of its own reflection." (Donald 2001, p326)

6 References

Blackmore, S. J. (1999), *The Meme Machine*, Oxford University Press.

Clark, A. and Chalmers, D. (1998), 'The extended mind', *Analysis*, **58**:7-19.

Clark, A. (2005), 'Intrinsic content, active memory and the extended mind', *Analysis*, **65**:1:1-11.

Clark, A. (to appear), 'Memento's Revenge: Objections and Replies to the Extended Mind' in R. Menary (ed) *Papers On The Extended Mind*.

Dawkins, R. (1976), *The Selfish Gene*, Oxford University Press.

Dennett, D. (1987), *The Intentional Stance*, MIT Press.

Logan, R. K. (2007), *The Extended Mind: The Emergence of Language, the Human Mind and Culture*, University of Toronto Press.

Rowlands, M. (2003), *Externalism: Putting Mind and World Back Together Again*, McGill-Queen's University Press.

Rupert, R. (2004), 'Challenges to the Hypothesis of Extended Cognition', *Journal of Philosophy* **101**: 389-428.

Searle, J. (1980), 'Minds, Brains, and Programs.' *Behavioral and Brain Sciences* **3**:417-424.

Wheeler (draft), 'Extended X: Recarving the Biological and Cognitive joints of Nature'.