MICHAEL HAUSKELLER

Department of Sociology and Philosophy, University of Exeter, Amory Building, Rennes Drive,
Exeter EX4 4RJ, UK
m.hauskeller@exeter.ac.uk

The progressing cyborgization of the human body reaches its completion point when the entire body can be replaced by uploading individual minds to a less vulnerable and limited substrate, thus achieving "digital immortality" for the uploaded self. The paper questions the philosophical assumptions that are being made when mind-uploading is thought a realistic possibility. I will argue that we have little reason to suppose that an exact functional copy of the brain will actually produce similar phenomenological effects (if any at all), and even less reason to believe that the uploaded mind, even if similar, will be the same self as the one on whose brain it was modeled.

Key words: mind-uploading, cyborgization, transhumanism, human enhancement, personal identity, functionalism

1. Introduction

We humans have always remodeled the external world according to our needs and desires. That is, of course, not very unusual. Most animals take active part in rebuilding their environment to construct a suitable niche for themselves [Odling-Smee, 2003]. Yet no animal does it so thoroughly and extensively as we do. We have never really stopped being busy making our environment fit for us, have never tired of constantly rebuilding the world in such a way that it assists us in our will to live, and to live well. That is part of what makes us human. However, when it comes to our survival and well-being, a potentially hostile environment is only part of the problem. The world we live in may, if unchecked, thwart our aspirations and even kill us, but it is our own human body that allows this to happen in the first place. So in order to be safe, it seems that controlling our environment is not enough; we also need to gain complete control over our bodies to compensate for their natural frailty. But just as controlling the external world largely means replacing a natural environment by an artificial, human-produced one (retaining only those aspects of the former that are found useful, or harmless), controlling the human body also means replacing those parts of it that can no longer perform their function, that foreseeably will one day fail in performing it, or that do not perform their function as well as we may wish. This replacement is as much therapy as it is enhancement. Given that the human body, as it is naturally constituted, is *itself* a danger to our continued well-being, any improvement on the body's natural constitution is a remediation of a defect and thus therapeutic.

The replacement of human body parts by devices that simulate their function is of course not new. Artificial limbs have been used for thousands of years. Today the technology is so advanced that limbs can be controlled directly through conscious thought, which initiates muscle contractions that are converted into electronic signals, which in turn move the limb. Even though it appears that users of such advanced devices are still not entirely happy with the results, suggesting that as yet prosthetic limbs cannot adequately replace biological limbs [Biddiss & Chau, 2007], this may soon be the case.

Artificial cardiac pacemakers implanted in the body have been successfully used for fifty years now, and artificial hearts and lungs are already being tested on live patients. Cochlear implants can replace ears and compensate for loss of hearing by stimulating nerve fibers directly in the brain. Similar devices are being developed for the restoration of sight. By connecting a video camera to a blind person's brain that sends signals directly into their visual cortex, William Dobelle and his teams managed to restore their ability to see the outlines of objects and thus to use their sight to navigate their bodies [Dobelle, 2000]. Another device, called *Eyeborg*, invented by Adam Montandon in collaboration with the colour-blind artist Neil Harbisson translates visual data into sound-waves, which are then interpreted (though not directly *seen*) as colors [Wade, 2005].

All these developments indicate that, in principle, our sense organs are dispensable since we can create their effects, or more precisely what they allow us to know and do, by other means. We don't even seem to need a body, at least not a functioning one, to interact with the world. *BrainGate*, a brain implant developed by the American company *Cyberkinetics* in 2004 (now *BrainGate Co.*), allows patients that, due to a spinal cord injury, suffer a total sensory and motor loss of their limbs and torso to control external devices connected to a computer through muscle contractions initiated by their thoughts [Hochberg, 2006]. Even the brain itself might be replaceable. Alzheimer patients and others suffering from the effects of brain damage located in the hippocampus may soon be helped by having the damaged parts replaced by an artificial brain prosthesis, a microchip hippocampus [Berger, 2005; Berger & Glanzman, 2005].

2. Messy Bodies

What we witness here is what is often described as an increasing cyborgization of the human, where ‘cyborg’ can be defined as a human being some of whose parts are artificial. In light of these developments it may appear not unreasonable to expect that this is only the beginning and we will progress further until we have achieved the goal that is implicitly pursued in all those innovations that couple human beings with fast-paced hypertechnology: complete independence from nature, unrestricted autonomy. For as long as we are hooked to this organic body, we will never be entirely free and safe. The organic body is a limitation that is resented by many, and that they hope we’ll be able to overcome not too far in the future. “Soon we could be meshing our brains to computers, living, for all practical purposes, on an ‘immortal’ substrate, perhaps eventually discarding our messy, aging, flesh-and-bones body altogether.” [Klein, 2003] The human body is not only regarded as dispensable; it is an obstacle, an enemy to be fought and to get rid of. It ages and makes us age with it, eventually annihilating us. It is “messy”, disorderly and dirty; it brings chaos and decay into our lives. “Flesh-and-bones” is a material that is deemed unsuitable for an advanced, dignified, enlightened and happy existence. So let’s abandon it if we can. Good riddance to bad rubbish! “If humans can merge their minds with computers, why would they not discard the human form and become an immortal being?” [Paul & Cox, 1996: 21].

Yet in order to become truly immortal, our goal should be to become a ‘cyberbeing’, a being that is more than just interlinked with machines, more than just partly a machine itself, and even more than a machine in its entirety. Gradually replacing human biology and the messy organic body by a more durable and more controllable substrate is certainly a considerable improvement, but is by no means sufficient. Why not go a step further and, if at all possible, discard the physical body altogether? That is, any *particular* body, any body that is *essentially* and not merely accidentally ours, not only something we use and can discard when proved not useful enough or no longer useful, but rather something that defines our very existence and has, as it were, pretensions of *being* us. In other words, why not relocate and transform our existence in such a way that we are no longer bound to any particular material substrate, be it organic or non-organic, because all we need, if anything at all, is the occasional body to-go as a communication facilitator, a hardware on which to run the program which we then will be [Moravec, 1989]. “Imagine yourself a virtual living being with senses, emotions, and a consciousness that makes our current human form seem a dim state of antiquated existence. Of being free, always free, of physical pain, able to repair any damage and with a downloaded mind that never dies.” [Paul & Cox, 1996: xv] The *telos*, the logical end point, of the ongoing cyborgization of the human is thus the attainment of “digital immortality”, which is more than just “a radical new form of human enhancement” [Sandberg & Bostrom, 2008: 5]. Rather, the desire to conquer death, that “greatest evil” [More, 1990], is its secret heart, that which gives the demands for radical human enhancement their moral urgency. And the best chance to attain what we desire is through the as yet still theoretical possibility of mind-uploading.

3. Minds are what Brains Do

“Uploading is the transfer of the brain’s mindpattern onto a different substrate (such as an advanced computer) which better facilitates said entity’s ends.” [Kadmon, 2003] To upload our minds to a computer would allow us not only to transfer our existence to a more durable substrate; it would allow us to roam the world of cyberspace without any clearly marked physical constraints or time limits. We could be anywhere and everywhere, all in a blink of an eye and for all time, until the world itself ends. How is this supposed to work? Theoretically, we will first scan the structure of a particular brain and then “construct a software model of it that (…), when run on appropriate hardware, will behave in essentially the same way as the original brain”, that is “produce the phenomenological effects of a mind”, or more precisely of a particular mind [Sandberg & Bostrom, 2008: 7]. At least that’s the idea. Whether or not that will really one day be possible nobody really knows, despite occasional protestations to the contrary. The scanning of a whole human brain is no doubt conceivable. So is the construction of a sufficiently accurate functional model of it. All we need for it is the right technology, and the exponentially growing speed with which technology has developed in the recent past certainly suggests that it won't be long before we have that technology available. It might, however, still prove too difficult a task to emulate a whole brain. *Moore's Law* or Kurzweil's *Law of Accelerating Returns* [Kurzweil, 1999: 30-33], according to which the time period between salient events (both in natural and in technological evolution) grows shorter the further we progress, might prove not to be laws at all but merely generalizations that describe fairly accurately what has happened in the past, but not necessarily what is going to happen in the future. Yet even if we will manage to emulate a whole brain, we may still find that the hoped for effect, namely that the model actually gives rise to subjective awareness, will fail to appear.

Whether or not "the phenomenological effects of a mind" will indeed appear obviously depends on what the mind is and in what way it is dependent in its existence on the body. We would have to assume that some form of functionalism is true, that, in the words of AI pioneer Marvin Minsky, minds are what brains *do* [Minksy, 1986: 287]. This generally means that the mind supervenes on the physical and is not dependent on the specific material constitution of what it supervenes on. We are to assume that the mind is based on the functional relations between physical elements and not on those elements themselves. Most versions of functionalism allow for multiple realizability, which means that they are open to the possibility that mind, though not necessarily the *same* mind, can be implemented by different physical properties. Hence we should be able to recreate a particular mind, say mine, by any means that permit the recreation of my mind's (or more precisely, my brain's) functional relations. So, as Ned Block has pointed out, if we were able to assign binary values to each member of the Chinese population and then persuade them to simulate those relations by following strict instructions of input-output regulation (i.e., of what to do in response to what they perceive is happening), then, if functionalism is true, that should result in the emergence of a mind such as mine [Block, 1978].

While this sure looks like an absurd consequence because we find it hard to imagine that people could be used to create such a thing as a mind, it is not exactly a convincing refutation of functionalism. Since nobody has yet, for obvious reasons, attempted to organize a large population of people to simulate a neural network and thus create a mind, we cannot completely rule out the possibility that it may actually work. Whether the mind really can be disconnected from its current biological substrate or whether it is dependent on the special causal powers of the organic brain, as John Searle has argued [Searle, 1990], is a matter of mere speculation, as long as we haven't had the chance to put the theory to the test by actually producing an accurate whole brain emulation and then seeing what happens. In other words, it is an empirical question, which we cannot decide on purely philosophical grounds.

However, the artificial creation of *a* mind is one thing, the recreation or transplantation of a *particular* mind that actually belongs to someone as *their* mind quite another. If what we want to achieve is not merely the creation of an intelligent, that is actually conscious and self-conscious, *machine* or the creation of a *model* of the human mind in order to improve our understanding of how it works, but rather some form of personal *immortality*, then we will have to make sure that the mind that does appear when we simulate the functional relations of a particular brain is not only qualitatively identical to the mind it is modeled on, but also numerically identical to it. Yet even qualitative identity is far from certain when a particular brain has been successfully emulated. It is entirely conceivable (though perhaps unlikely) that an artificial brain or brain substitute A* that mirrors accurately all functional relations of a particular organic brain A, although indeed producing certain phenomenological effects, does not produce the *same* effects as the original, organic brain. This is just assumed. But as first Jorge Luis Borges [1964] and then Arthur C. Danto [1981] have shown with their thought experiments, identical syntax does not guarantee identical semantics. Two artworks can be absolutely indistinguishable as objects and still be entirely different as artworks, not because they are interpreted in a different way, but because they are simply not *about* the same thing. The *form* is identical, but the objective (!) *meaning* is different. Similarly, we may find that the actual phenomenological constitution of the mind is not as thoroughly causally determined by the structure of the brain as we like to assume.

But even if the mind resulting from a particular brain emulation will indeed be qualitatively identical to the mind it is supposed to copy, the two minds also need to be *numerically* identical, which is an *additional* requirement. This means that for the successfully instantiated mind to be mine, it would *not* be sufficient if it were in every respect indistinguishable from my mind. Rather, it would have to be literally the same. There seems to be a conceptual difference between a mind that actually is mine, and a mind that merely thinks, feels, and remembers exactly as I do, and hence is *like* me in all respects, except that it happens not to *be* me. Think of two copies of the same book that, despite their having exactly the same design and content, are still *two* copies and not identical to each other.

4. Is the Mind like a Story in a Book?

But hang on, isn't that exactly the point of the functionalist theory of mind that the *same* mind can be instantiated by two or more material substrates that may differ considerably from each other in their make-up and appearance, as long as they are functionally equivalent? Let's have another look at the book analogy. It is true that two copies of the same book are still *two* copies. But it seems also to be true that these two copies are copies of the *same* book, that is to say, the same literary (fictional or non-fictional) entity that is represented by the symbols we find in both of them. Say you have two different editions of Joyce's *Ulysses*. Both editions are very different in appearance. The covers are different, as is the paper used, the page size, the size and typeface of the letters, and so on. In fact, there's hardly anything that is the same in both editions, except, that is, the story itself. The *Ulysses* as a particular literary creation seems to be equally present in both, so that if one of the two copies were destroyed, the *Ulysses* itself would easily survive the destruction. And we can imagine an indefinite number of editions of the *Ulysses*, all different from each other, but all containing, or instantiating, exactly the same story. For that story to exist and to continue to exist, it does not even have to be actually printed. It might only be available in electronic form as an e-book to be read at a computer screen, or as an audio-book to be listened to rather than read. Or it might only exist in the mind of a single person, as in Ray Bradbury's novel *Fahrenheit 451*, where various books, that is, specific literary entities, are saved from oblivion by having book-keepers assigned to them whose task it is to memorize and thus preserve them for future generations.

Can we not understand the mind in a similar fashion, as a distinct (though evolving) mental entity that can be indefinitely replicated in various forms without thereby changing its identity? If the mind is rather like a book and the brain like the specific material representation of that book (i.e., the book as a concrete material object), then there seems to be no question that a mind that is syntactically and semantically identical to mine, really *is* mine, just like the *Ulysses* is literally the *same Ulysses* (and not a *further Ulysses*) in each of its various representations.

However, there's one element here that is forgotten in this analogy. This element is the reader. Without the reader (or listener, or thinker) the *Ulysses* does not exist. It is instantiated primarily not in a material object but in a mind that interprets a certain series of symbols in a certain way and thereby creates or recreates the specific entity known as the *Ulysses*.

Now let us suppose that you are reading a copy of that particular book and I am reading, at the very same time and at the same pace, a copy of the same book. Let us further suppose that we are both utterly immersed in it and do not think of anything else but what we are reading. One might then say that our minds, while we are reading, are qualitatively identical with each other. However, it seems that you would still be you and I would still be I, and not simply because we happen to inhabit different bodies. If our minds are distinct entities in the first place then there is no reason they should not remain distinct even when their contents happen to become identical. It seems strange to assume that as long as you and I have different thoughts (and I'm using the word 'thought' here in the wide sense of Descartes' 'cogito') we are distinct entities, but as soon as we think alike we become one and the same. Just like a book, in order to exist, needs a reader, the mind needs a *self*. And just as the reader is not the book, the self is not the mind. Rather, it is a particular *appropriation* of the mind. The same book can be read by different readers. Similarly, the same mind might be 'read' or had by different selves. Note that I am not claiming that this is actually possible. I am just saying that it is *conceivable*.

However, it has been argued that the notion of different selves makes no sense, because 'self' is just a particular quality of the mind, which is the same quality in all minds: "An experience must be a universal across times as well as across brains. This experience of being you, here *now*, would be numerically the same *whenever*, as well as wherever, it was realized" **[**Zuboff, 1990**]**. We can call this quality 'immediacy'. If this view is accepted then my above argument fails. Identical minds could not be had by different selves. So if I managed to recreate my mind by uploading its entire content to a computer, I would also have succeeded in recreating my self. However, if there is in fact only one self, then there is also only one self when the minds are *different* from each other, as it is commonly the case. In other words, you are I and I am you even if our minds are not at all alike. But if that is so then mind-uploading becomes needless, because I'm already immortal. For when I, that is, this particular instantiation of the self, dies then I continue to live in you, as in fact I have been doing all along.

5. Still Jack?

So how likely is it that a software model of my mind, uploaded to a computer, will really be *my* mind in the sense that it is really *I* that will exist in this new material form? Ray Kurzweil has brought forward an argument that seems to show that this is most likely indeed [Kurzweil, 1999: 52-55]. He asks us to imagine a person ("Jack") who starts out as a normal human being and then gradually has parts of his body replaced by better, artificial ones. He begins with cochlear implants and then ads, step by step, other devices, advanced imaging processing implants, memory implants, and so on, until, eventually, he takes the final step of replacing his whole brain and neural system with electronic circuits. In other words, he has his brain scanned and the information then stored in a computer. Is he, after the completion of that final stage, still the same person, still the same old Jack? Well, we probably would want to say that he is still the same person after he has had his cochlear implant. And also with enhanced vision and enhanced memory we would hesitate to call him a different person. "Clearly", writes Kurzweil, "he has changed in some ways and his friends are impressed with his improved faculties. But he has the same self-deprecating humor, the same silly grin – yes, it's still the same guy." So where exactly do we draw the line? Where does Jack cease to be Jack? It seems that if Jack remains the same person after each replacement, there is no reason he should suddenly become a different person (in the sense of ceasing to exist and being replaced by someone else) after the last and final replacement. Let's call this the *argument from graduality*. According to Kurzweil it shouldn't make any difference if Jack's body is replaced gradually or in one go. Even if the transition from an entirely organic body to a machine takes place in one quick step, Jack will still be Jack. But will he really?

Well, he may still have his "self-deprecating humor", but his "silly grin" surely has been lost in the transition. Does that matter? To Jack's friends it might, but it is irrelevant to the question we are trying to answer - as is, by the way, Jack's sense of humor. Because even if Jack changes completely and loses all the little peculiarities by which his friends used to recognize him, so that they might say that he has become a different person altogether, it might still be him in the sense of being the same self. Over the course of a life time I may change considerably, but as long as *I* undergo this change I am the same self - though not necessarily the same person, depending on how we want to define the word 'person'. If we take 'person' to mean for example with John Locke "a thinking intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places" [Locke, 1753: II. Xxvii.9], then it is not entirely clear that I am still the same person I used to be when I was a child of two. But the same self I surely am, because *what* I am is one thing, *that* I am quite another. However, although I remain the same self, that is the same subject of experience, despite many changes in my character, the converse is also true, and for the same reasons: I might not seem to change at all (same "self-deprecatory humor" and all), but still cease to be, while a different self takes my place that has all the properties I, or my mind, used to have, without *being* me at all. So again, will Jack still be Jack after he has given up his organic body for a life as a program installed in a computer? That is what the argument from graduality suggests, but on what grounds exactly?

6. Graduality and Identity

Kurzweil's argument is clearly a variation of the ancient paradox known as the *Ship of Theseus*, in which we are asked to imagine a ship that is being maintained by having those of its parts replaced that are no longer functional, until eventually all of its original parts are gone and everything is new. The question then arises whether the ship is still the same ship or rather a different one. Some will argue that since there is no step along the way of gradual material change where we can say that *now* the ship is a different one, it remains the same all the way through, whereas others will argue that it cannot be the same since the material constituents which together formed the ship no longer exist (or if they do exist, perhaps used to build a new ship, then this ship would be identical to the original one rather than the one with the replaced parts). Who is right? Well, neither. It is a paradox precisely because there is no obvious answer to the question. And that is because a ship might have a name, but it does not have an identity, not in and by itself. And because it doesn't have one, it cannot change it suddenly. *We* preside over its identity, and it makes no difference whatsoever to the *ship* whether or not we call it the same after it had all its parts replaced. In certain contexts and for certain purposes it might be appropriate to call it the same ship, and in other contexts and for other purposes it might be more appropriate to call it a different ship. Unfortunately, we don't have that option with people. A person either is the same self or they are not, and whether we *regard* them as the same or as different has got nothing to do with it.

The argument from graduality is in fact *always* fallacious because it denies the reality of change. It seems reasonable to assume that a heap of sand from which one single grain is removed will still be a heap, but even though we are not able to pinpoint the exact moment when, after the removal of yet another grain, the heap ceases to be a heap (because there is no such moment) we will at some point all agree that it is no longer one (*Sorites* paradox). At some stage (and it might not always be the same) we will begin to doubt whether 'that thing there' is still a heap. From then on our reluctance to still call it so will grow until it has sufficiently diverged from our idea of a heap that it will no longer seem appropriate to call it that. Similarly, when a person gradually changes, so that their character eventually is very different from what it used to be, we will at some stage begin to doubt whether it is the still the same person until eventually we will accept that the person we used to know, in a manner of speaking, no longer exists. But there's no definite threshold here, no point in time where one becomes another person, or rather character. However, there are distinctive and radical changes that really *are* sudden. If you applied the argument from graduality to, say, the states of matter by reasoning that since water of a temperature of 20 degrees Celsius is liquid, it will still be liquid if we lower its temperature by one degree, you will obviously be proven wrong, provided you repeat the procedure often enough, because we all know that the water will start turning to solid ice at some clearly defined point. That is no gradual change and it has got nothing to do with the vagueness of our concepts. Rather, the change is comparatively sudden and very real. And it might well be that when it comes to the individual self we are dealing with a similar situation. The self may survive the gradual cyborgization of the human body through various changes, but only up to a certain critical point. When that point is reached the self might disappear (to be either replaced by a different self or to vanish entirely without being replaced at all).

7. The Human Self

It is worth pointing out, though, that while, for all we know, the identity of the *self* may change abruptly, our identity as *human beings* is not likely to do so because it is not confined by sharply demarcated boundaries. This means that my self is not necessarily a human self. Although we may find ourselves reluctant to call a being human that exists as a software program in a replaceable robotic body, this does not imply that the no-longer-human is not the same individual they used to be when they were still human. Just as I can remain myself even when my character changes so considerably that my friends are no longer able to recognize me, it seems that I can still be I after I have shed my humanity. One doesn't lose one's humanity like one loses one's virginity, or one's job, or one's self: in one decisive step. Taken by itself losing one's humanity is not a real change at all. The predicate 'human' is itself a human classification, just as the word 'heap' is, so that my continuing humanity depends on the elasticity of the currently prevalent image (or images) of the human [Hauskeller, 2009). Only after what I have become can no longer be aligned with the vague and changing idea that people have about themselves as humans, I am human no longer. Bostrom's 'Golden', the fictional retriever whose mind has been uploaded and cognitively enhanced so that he can now tell his story to a TV audience on the Larry King show can hardly be taken seriously as a dog [Bostrom, 2004]. He has become almost human, or something that is neither human nor dog. But he may still be the same individual he was before the upload. Or maybe not, if it turns out that one's self is *in fact* inseparable from the organic substrate in which it has developed.

8. Walking Algorithms?

The possibility of replacing one's brain and body presupposes an ontological distinction between body and mind, a particular kind of Cartesian substance dualism. The mind is supposed to be not the brain, but "structure" and "pattern", which contains "information" that can in principle always be separated from its organic basis, replicated and re-instantiated in an indefinite number of different material forms. The brain is the (replaceable) hardware and the mind the software – the ghost in the machine [Potts 1996]. However, this is not thought to be a peculiarity of the brain-mind relationship, but rather a particular case of a general feature of reality. Not only the mind is said to be nothing but information, but also the quality of being alive: "from the perspective of many contemporary biologists, life is just an interesting configuration of information." [Doyle, 2003: 20] "All living organisms are no more than walking algorithms." [Kadmon, 2003] Yet non-living things are also algorithms, except that they don't walk. They, too, are configurations of information, only less interesting ones. "Whatever is happening in the universe (…) is all information." [Paul & Cox 1996, 34] Thus the whole of reality is understood as being *essentially* information. Whatever else exists or seems to exist are just ways this information is conveyed and processed. It is the form of its appearance, but not the thing-in-itself. It is not the 'really real'. Just as Richard Dawkins once described living organism as vehicles for the replication of "selfish genes" [Dawkins, 1976], they are now being understood as vehicles for the preservation and transmission of information. As Katherine Hayles has pointed out correctly: "Underlying the idea of cyberspace is a fundamental shift in the premises of what constitutes reality. Reality is considered to be formed not primarily from matter or energy but from information. Although information can be carried by matter or energy, it remains distinct from them. Properly speaking, it is a pattern rather than a presence." [Hayles, 1996: 112] However, the truth of this *information idealism* is far from obvious. Rather, it is a substantial metaphysical claim, which ought to be treated as such. We need to ask whether the assumption is at all justified. Are we really no more than bits of information? Are we just "walking algorithms"?

It seems to me that not even the mind, let alone a particular person, can be adequately described as information if what we mean by the word 'mind' is more than just content. And we usually do. Our own minds at least are conscious, and it is this consciousness that seems to make them minds in the first place. It is doubtful whether there can be minds that are not conscious (without stretching the meaning of the term beyond recognition). Having a mind generally means being to some extent *aware* of the world and oneself, and this awareness is not itself information. Rather, it is a particular way in which information is processed (which is different from the way in which, say, information in an electronic circuit is processed), but this way doesn't add anything to the already existing information. It is not simply information about how to process other bits of information. That is why, theoretically, your mind can contain the same information as mine and still not *be* mine (or, as in split-brain patients, different information and be mine nonetheless). Even though we don't understand how it is possible for there to be such a separation, your mind will always stay yours, and mine, mine. For the mind is always *somebody's* mind, which means that there can be a mind only when there is a *self*. A mind without a self is inconceivable. What is conceivable is a mind that is qualitatively identical to mine and yet somebody else's.

9. Or World-involving Beings?

However, that there is no mind without a self does not imply that a particular self is nothing but a particular mind. This, too, is an unwarranted Cartesian assumption. Though my mind is part of what I am, I am not my mind. I am there even when I'm completely unconscious. I breathe and the blood is pumping through my veins (assisted perhaps by an artificial heart). I am, though my mind may be blank. My body harbors my self while my mind is absent. So I am not my mind. Neither am I my brain, though it is no doubt also a part of me. We tend to exaggerate the importance of the brain to the extent that we confuse our own actions with the actions of the brain. Statements such as the following we find plenty: "The human brain is one that loves, feels empathy, projects into the future, and contemplates a lifetime of memories. It is subject to pleasure and joy, and a good laugh." [Paul & Cox, 1996: 273] The brain does, of course, none of these things. It doesn't laugh, and neither does it love, feel empathy etc. *We* do. The brain certainly helps but it is still we who act and relate to our environment. The brain is only one of our organs (albeit a very important one), that is, an instrument that we use in order to accomplish certain tasks in accordance with our general desire to survive in this world. My brain is situated in a body, as is my mind, which is one of my modes of existence, no more and no less. Although, let's face it, we don't have the slightest clue how conscious experience comes about and how there can be such things as selves in the first place, it is rather unlikely that mind and self are directly produced by the brain, as is commonly assumed. There is no direct evidence for that. The brain develops and changes with the experience we accumulate during our lives, and it does so because it has a particular job to do within the system that we call a living, conscious being. It rises to the occasion. That we can manipulate the mind by manipulating the brain, and that damages to our brains tend to inhibit the normal functioning of our minds, does not show that the mind is a product of what the brain does. The brain could be just a facilitator. When we look through a window and the window is then painted black, our vision is destroyed or prevented, but we cannot infer from this that the window produces our ability to see. The brain might be like a window to the mind. Surely the mind is not in any clear sense localized in the brain. Alva Noe is right when he declares the locus of consciousness to be "the dynamic life of the whole, environmentally plugged-in person or animal" [Noe, 2009: xiii] We are not our brains, we are "out of our heads", as Noe puts it, reaching out to the world as "distributed, dynamically spread-out, world-involving beings." [Noe, 2009: 82]

The only selves we have ever encountered are situated, embodied selves, agents that interact with the world and each other through and in their bodies and minds. Even ourselves have we never known in any other way. And when we relate to other selves we always relate to them as complex wholes that do not merely exist as minds (and whose apparent location in a particular body is purely accidental).

When we love someone we love them for what they are, and what they are is not hidden away. Rather, what they are, is manifest in their bodies, in the "way you hold your knife", the "way you sip your tea", and yes, in Jack's "self-deprecatory smile" too. People are there for us "body and soul". It would be hard, if not downright impossible, to love (or, for that matter, to hate, or care in any way for) a software program, even if it were conscious. Perhaps we couldn't even care for ourselves then. So if we really managed one day to upload our minds to a computer the best we could achieve thereby would be the continuation of a stripped-down, rudimentary self. And that is probably not the kind of immortality most of us would regard as desirable.

10. Conclusion

I have argued that the hope of attaining "digital immortality" through a completion of the ongoing cyborgization process, i.e. through mind-uploading, rests on several questionable assumptions. We have no evidence whatsoever to support the idea that (1) even a perfectly accurate software emulation of a human brain will actually result in conscious experience. This will only happen if the formalist theory of mind is true, which we have no way of knowing and which we have no reason to believe until we encounter a mind instantiated in something that is not a living, organic body. But if it does result in conscious experience, then we (2) still have no guarantee that it will be anything like the experience of the mind we intended to duplicate, or recreate. Should this be the case, though, then we may (3) still see our hopes disappointed by the fact that the newly created mind, though qualitatively identical to the mind whose continued existence we intended to ensure through the emulation, is not numerically identical to it. In other words, it may be a different mind, or more precisely, a different self. Although the self may be preserved through various stages of increasing cyborgization, the final step may prove one step too far and end the existence of the self. This is not only possible, but indeed very likely, since the final step is different from the previous ones in so far as it relies on the possibility of *copying* the self (instead of merely preserving it through a series of changes). Yet the only thing that *can* be copied is information, and the self, *qua* self, is not information. But even we managed to not lose the self during the copying process and to somehow connect it to the new non-organic substrate, we would (4) have trouble recognizing ourselves. For what we think of as ourselves is very much tied to our bodily existence and as such far more comprehensive and richer than a mere mind can ever be.

References

Berger.T.W., et al. [2005] Restoring Lost Cognitive Function. Hippocampal–Cortical Neural Prostheses. *IEEE, Engineering in Medicine and Biology Magazine* 24(5), 30-44.

Berger, T.W., and Glanzman, D.L. (eds.) [2005] *Toward Replacement Parts for the Brain: Implantable Biomimetic Electronics as Neural Prostheses* (MIT Press).

Biddiss, E.A., & Chau, T.T. [2007] Upper limb prosthesis use and abandonment: A survey of the last 25 years, *Prosthetics and Orthotics International* 31(3), 236-257.

Block, N. [1978] Troubles with Functionalism, *Minnesota Studies in the Philosophy of Science* 9, 261-325.

Borges, J.L. [1964] Pierre Menard, Author of the Quixote, in Borges, *Labyrinths: Selected Stories and Other Writings* (New Directions Books, New York), 36-44.

Bostrom, N. [2004] Golden, .www.nickbostrom.com.

Danto, A.C. [1981] *The Transfiguration of the Commonplace* (Harvard University Press, Cambridge, MA).

Dawkins, Richard [1976] *The Selfish Gene* (Oxford University Press, New York).

Dobelle, W.H. [2000] Artificial Vision for the Blind by Connecting a Television Camera to the Visual Cortex, *ASAIO Journal* 46, 3-9.

Doyle, R. [2003] *Wetwares*. *Experiments in Postvital Living* (University of Minnesota Press, Minneapolis).

Hauskeller, M. [2009] Making Sense of What We Are, *Philosophy* 84(1), 95-109.

Hayles, N.K. [1996] How Cyberspace Signifies: Taking Immortality Literally, in G. Slusser et al. (eds.), *Immortal Engines. Life Extension and Immortality in Science Fiction and Fantasy* (University of Georgia Press, Athens and London), 111-121.

Hochberg, L.R., & al. [2006] Neuronal ensemble control of prosthetic devices by a human with tetraplegia, *Nature* 442, 164-171.

Kadmon, A. [2003] Transtopia – Transhumanism Evolved, http://www.transtopia.org/uploading.html.

Kurzweil, R. [1999] *The Age of Spiritual Machines* (Viking Penguin, New York).

Locke, J. [1753] *An Essay Concerning Human Understanding* (14th ed. S. Birt et al., London).

Klein, B. [2003] Building a Bridge to the Brain, http://www.imminst.org/forum/index.php?act=ST&f=67&t=938&s.

Minsky, M. [1986] *The Society of Mind* (Simon and Schuster, New York).

More, M. [1990] Transhumanism. A Futurist Philosophy, *Extropy* 6, 6-12.

Moravec, H. [1989] *Mind Children: The Future of Robot and Human Intelligence* (Harvard University Press, Cambridge, MA).

Noe, A. [2009] *Out of Our Heads* (Hill and Wang, New York).

Odling-Smee, P., et al. [2003] *Niche Construction: The Neglected Process in Evolution* (Princeton University Press, Princeton).

Paul, G.S., & Cox, E.D. [1996] *Beyond Humanity: CyberEvolution and Future Minds* (Charles River Media, Rockland, MA).

Potts, St. [1996] IBMortality: Putting the Ghost in the Machine, in G. Slusser et al. (eds.), *Immortal Engines. Life Extension and Immortality in Science Fiction and Fantasy* (University of Georgia Press, Athens and London), 102-110.

Sandberg, A. & Bostrom, N. [2008] *Whole Brain Emulation: A Roadmap* (Future of Humanity Institute, Oxford).

Searle, J. [1990] Is the Brain's Mind a Computer Program?, *Scientific American* 262, 26-1.

Wade, G. [2005] Seeing things in a different light, http://www.bbc.co.uk/devon/news_features/2005/eyeborg.shtml.

Zuboff, A. [1990] One Self. The Logic of Experience, *Inquiry* 30, 39-68.