19

MINDSCAN: TRANSCENDING AND ENHANCING THE HUMAN BRAIN[1]

Susan Schneider

Suppose it is 2025 and, being a technophile, you purchase brain enhancements as they become readily available. First, you add a mobile internet connection to your retina, then you enhance your working memory by adding neural circuitry. You are now officially a cyborg. Now skip ahead to 2040. Through nanotechnological therapies and enhancements you are able to extend your lifespan, and as the years progress, you continue to accumulate more far-reaching enhancements. By 2060, after several small but cumulatively profound alterations, you are a "posthuman." To quote philosopher Nick Bostrom, posthumans are possible future beings, "whose basic capacities so radically exceed those of present humans as to be no longer unambiguously human by our current standards" (Bostrom 2003). At this point, your intelligence is enhanced not just in terms of speed of mental processing; you are now able to make rich connections that you were not able to make before. Unenhanced humans, or "naturals," seem to you to be intellectually disabled – you have little in common with them – but as a transhumanist, you are supportive of their right to not enhance (Bostrom 2003; Garreau 2005; Kurzweil 2005).

It is now AD 2400. For years, worldwide technological developments, including your own enhancements, have been facilitated by superintelligent AI. A superintelligence is a creature with the capacity to radically outperform the best human brains in practically every field, including scientific creativity, general wisdom, and social skills. Indeed, as Bostrom explains, "creating superintelligence may be the last invention that humans will ever need to make, since superintelligences could themselves take care of further scientific and technological developments" (Bostrom 2003). Over time, the slow addition of better and better neural circuitry has left no real intellectual difference in kind between you and a superintelligent AI. The only real difference between you and an AI creature of standard design is one of origin – you were once a "natural." But you are now almost entirely engineered by technology – you are perhaps more aptly characterized as a member of a rather heterogeneous class of AI life forms (Kurzweil 2005).

So let me ask: should you enhance and if so, why? I have just given a very rough sketch of the kind of developmental trajectory that the transhumanist generally aspires to.[2] Transhumanism is a philosophical, cultural, and political movement that holds that the human species is now in a comparatively early phase and that its very evolution will be altered by developing technologies. Future humans will be very unlike their present-day incarnation in both physical and mental respects, and will in fact resemble certain persons depicted in science fiction stories. Transhumanists share the belief that an outcome in which humans have radically advanced intelligence, near immortality, deep friendships with AI creatures, and elective body characteristics is a very desirable end, both for one's own personal development and for the development of our species as a whole.

Despite its science fiction-like flavor, the future that transhumanism depicts is very possible: indeed, the beginning stages of this radical alteration may well lie in certain technological developments that either are already here (if not generally available), or are accepted by many in the relevant scientific fields as being on their way (Roco and Bainbridge 2002; Garreau 2005). In the face of these technological developments, transhumanists offer a progressive bioethics agenda of increasing public import. They also present a thought-provoking and controversial position in philosophy of cognitive science, applying insights about the computational nature of the mind to the topic of the nature of persons, developing a novel version of one popular theory of personal identity: the psychological continuity theory.

In this chapter I shall employ science fiction thought experiments to discuss what I take to be the most important philosophical element of the transhumanist picture – its unique perspective on the nature and development of persons. Persons are traditionally viewed as being an important moral category, being the bearers of rights, or at least deserving of consideration of their interests in a utilitarian calculus. And, as we shall see, considering the nature of persons through the lens of transhumanism involves pushing up against the boundaries of the very notion of personhood. For consider again the issue of enhancement. When one wonders whether to enhance in the radical ways the transhumanists advocate, one must ask, "Will this radically enhanced creature still be me?" If not, then, on the reasonable assumption that one key factor in a decision to enhance oneself is one's own personal development, even the most progressive technophile will likely regard the enhancement in question as undesirable, for when you choose to enhance in these radical ways, the enhancement does not really enhance *you*. As we shall soon discuss, this is a lesson that the main character in Hugo award winner Robert Sawyer's *Mindscan* learns the hard way. Hence, examining the enhancement issue from the vantage point of the metaphysical problem of personal identity will thereby present a serious challenge to transhumanism. Given

their conception of the nature of a person, radical, and even mild, enhancements are risky, not clearly resulting in the preservation of one's original self. Indeed, I suspect that this is a pressing issue for any case for enhancement.

The Transhumanist Position

Transhumanism is by no means a monolithic ideology, but it does have an organization and an official declaration. The World Transhumanist Association is an international nonprofit organization that was founded in 1998 by philosophers Nick Bostrom and David Pearce. The main tenets of transhumanism were laid out in the Transhumanist Declaration (World Transhumanist Association 1998) and are reprinted below:

1. Humanity will be radically changed by technology in the future. We foresee the feasibility of redesigning the human condition, including such parameters as the inevitability of aging, limitations on human and artificial intellects, unchosen psychology, suffering, and our confinement to the planet earth.
2. Systematic research should be put into understanding these coming developments and their long-term consequences.
3. Transhumanists think that by being generally open and embracing of new technology we have a better chance of turning it to our advantage than if we try to ban or prohibit it.
4. Transhumanists advocate the moral right for those who so wish to use technology to extend their mental and physical (including reproductive) capacities and to improve their control over their own lives. We seek personal growth beyond our current biological limitations.
5. In planning for the future, it is mandatory to take into account the prospect of dramatic progress in technological capabilities. It would be tragic if the potential benefits failed to materialize because of technophobia and unnecessary prohibitions. On the other hand, it would also be tragic if intelligent life went extinct because of some disaster or war involving advanced technologies.
6. We need to create forums where people can rationally debate what needs to be done, and a social order where responsible decisions can be implemented.
7. Transhumanism advocates the well-being of all sentience (whether in artificial intellects, humans, posthumans, or non-human animals) and encompasses many principles of modern humanism. Transhumanism does not support any particular party, politician or political platform.

This document was followed by the much longer and extremely informative *The Transhumanist Frequently Asked Questions*, authored by Nick Bostrom, in consultation with dozens of leading transhumanists (Bostrom 2003).[3]

The Nature of Persons

Now let us consider some of the ideas expressed in the Declaration. Overall, central transhumanist texts have advanced a sort of trajectory for the personal development of a contemporary human, technology permitting (Kurzweil 1999, 2005; Bostrom 2003, 2005).[4]

> 21st century unenhanced human → significant "upgrading" with cognitive and other physical enhancements → posthuman status → "superintelligence"[5]

Recalling the chronology of enhancements I sketched at the beginning of this chapter, let us again ask: Should you embark upon this journey? Here, there are deep philosophical questions that have no easy answers.[6] For in order to understand whether you should enhance, you must first understand what you are to begin with. But what is a person? And, given your conception of a person, after such radical changes, would you yourself continue to exist, or would you have ceased to exist, having been replaced by someone else? If the latter is the case, why would you want to embark on the path to radical enhancement at all?

To make such a decision, one must understand the metaphysics of personal identity – that is, one must answer the question: What is it in virtue of which a particular self or person continues existing over time? A good place to begin is with the persistence of everyday objects over time. Consider the espresso machine in your favorite café. Suppose that five minutes have elapsed and the barista has turned the machine off. Imagine asking the barista if the machine is the same one that was there five minutes ago. She will likely tell you the answer is glaringly obvious – it is of course possible for one and the same machine to continue existing over time. This seems to be a reasonable case of persistence, even though at least one of the machine's properties has changed. On the other hand, if the machine disintegrated or melted, then the same machine would no longer exist. What remained wouldn't be an espresso machine at all, for that matter. So it seems that some changes cause a thing to cease to exist, while others do not. Philosophers call the characteristics that a thing must have as long as it exists "essential properties."

Now reconsider the transhumanist's trajectory for enhancement: for radical enhancement to be a worthwhile option for you, it has to represent a form of personal development. At bare minimum, even if enhancement brings such goodies as superhuman intelligence and radical life extension, it must not involve the elimination of any of your essential properties. For in that case, the sharper mind and fitter body would not be experienced by you – they would be experienced by someone else. Even if you would like to become superintelligent, knowingly embarking on a path that trades away one or more of your

essential properties would be tantamount to suicide – that is, to your intentionally causing yourself to cease to exist. So before you enhance, you had better get a handle on what your essential properties are.

Transhumanists have grappled with this issue. Ray Kurzweil asks: "So who am I? Since I am constantly changing, am I just a pattern? What if someone copies that pattern? Am I the original and/or the copy? Perhaps I am this stuff here – that is, the both ordered and chaotic collection of molecules that make up my body and brain" (Kurzweil 2005: 383). Kurzweil is here referring to two theories at center stage in the age-old philosophical debate about the nature of persons. The leading theories include the following:

1. The soul theory – your essence is your soul or mind, understood as a non-physical entity distinct from your body.
2. The psychological continuity theory – you are in its most general form essentially your memories and ability to reflect on yourself (Locke) and, you are your overall psychological configuration, what Kurzweil referred to as your "pattern."[7]
3. Materialism – you are essentially the material that you are made out of – what Kurzweil referred to as "the ordered and chaotic collection of molecules that make up my body and brain" (Kurzweil 2005: 383).
4. The no self view – the self is an illusion. The "I" is a grammatical fiction (Nietzsche). There are bundles of impressions but no underlying self (Hume). There is no survival because there is no person (Buddha).[8]

Upon reflection each of these views has its own implications about whether one should enhance. If you hold (1), then your decision to enhance depends on whether you believe the enhanced body would retain the same soul or immaterial mind.[9] If you believe (3), then any enhancements must not change your material substrate. In contrast, according to (2), enhancements can alter the material substrate but must preserve your psychological configuration. Finally, (4) contrasts sharply with (1)–(3). If you hold (4), then the survival of the person is not an issue, for there is no person to begin with. You may strive to enhance nonetheless, to the extent that you find intrinsic value in adding more super-intelligence to the universe – you might value life forms with higher forms of consciousness and wish that your "successor" should be such a creature.

Of all these views, (2) is currently the most influential, as philosopher Eric Olson underscores:

> Most philosophers believe that our identity through time consists in some sort of psychological continuity. You are, necessarily, that future being who in some sense inherits his mental features from you . . . the one who has the mental features he has then in large part because you have the mental features you have now. And you are that past being whose mental features you have inherited.

> . . . So magnetic is this view that many feel entitled to assert it without argument. (Olson 2002)

I will now suggest that the Transhumanist adopts a novel version of the psychological continuity view – namely, they adopt a computational account of continuity. First, consider that transhumanists generally adopt a computational theory of the mind.

> The Computational Theory of Mind ("CTM"): The mind is essentially the program running on the hardware of the brain, that is, the algorithm that the brain implements, something in principle discoverable by cognitive science.[10]

Computational theories of mind can appeal to various computational theories of the format of thought: connectionism, dynamical systems theory (in its computational guise), the symbolic or language of thought approach, or some combination thereof. These differences will not matter to our discussion.

In philosophy of mind, computational theories of mind are positions about the nature of thoughts and minds; unfortunately, discussions of CTMs in mainstream philosophy of mind do not generally speak to the topic of personhood. (Perhaps this is because personal identity is a traditional topic in metaphysics, not philosophy of mind.) But upon reflection, if you uphold a CTM, then, assuming you believe minds persist over time, it is quite natural to adopt a computational theory of persons. For note that proponents of CTMs reject the soul theory, for they reject the idea that minds are non-physical entities. One might suspect that the transhumanist views materialism favorably, the view which holds that minds are basically physical or material in nature and that mental features, such as the thought that espresso has a wonderful aroma, are ultimately just physical features of brains. Transhumanists reject materialism, however. For instance, consider Kurzweil's remark:

> The specific set of particles that my body and brain comprise are in fact completely different from the atoms and molecules that I comprised only a short while ago. We know that most of our cells are turned over in a matter of weeks, and even our neurons, which persist as distinct cells for a relatively long time, nonetheless change all of their constituent molecules within a month. . . . I am rather like the pattern that water makes in a stream as it rushes past the rocks in its path. The actual molecules of water change every millisecond, but the pattern persists for hours or even years. (Kurzweil 2005: 383)

Later in his discussion, Kurzweil calls his view "Patternism" (ibid.: 386). Put in the language of cognitive science, as the transhumanist surely would, what is essential to you is your computational configuration – for example, what

sensory systems/subsystems your brain has (e.g. early vision), the way that the basic sensory subsystems are integrated in the association areas, the neural circuitry making up your domain general reasoning, your attentional system, your memories, and so on – overall, the algorithm that your brain computes.[11]

Kurzweil's to patternism is highly typical of transhumanism. For instance, consider the appeal to patternism in the following passage of *The Transhumanist Frequently Asked Questions*, which discusses the process of uploading:

> Uploading (sometimes called "downloading," "mind uploading" or "brain reconstruction") is the process of transferring an intellect from a biological brain to a computer. One way of doing this might be by first scanning the synaptic structure of a particular brain and then implementing the same computations in an electronic medium. . . . An upload could have a virtual (simulated) body giving the same sensations and the same possibilities for interaction as a non-simulated body. . . . And uploads wouldn't have to be confined to virtual reality: they could interact with people on the outside and even rent robot bodies in order to work in or explore physical reality. . . . Advantages of being an upload would include: Uploads would not be subject to biological senescence. Back-up copies of uploads could be created regularly so that you could be re-booted if something bad happened. (Thus your lifespan would potentially be as long as the universe's.). . . . Radical cognitive enhancements would likely be easier to implement in an upload than in an organic brain. . . . A widely accepted position is that you survive so long as certain information patterns are conserved, such as your memories, values, attitudes, and emotional dispositions. . . . For the continuation of personhood, on this view, it matters little whether you are implemented on a silicon chip inside a computer or in that gray, cheesy lump inside your skull, assuming both implementations are conscious. (Bostrom 2003)

In sum, the transhumanist's cognitive science orientation introduces a new computationalist element to the traditional psychological continuity view of personhood. If plausible, this would be an important contribution to the age-old debate over the nature of persons. But is it correct? And further, is patternism even compatible with enhancement? In what follows, I suggest that patternism is deeply problematic. Furthermore, as things now stand, patternism is not even compatible with the enhancements that the transhumanists appeal to.

Robert Sawyer's *Mindscan* and the Reduplication Problem

Jake Sullivan has an inoperable brain tumor. Death could strike him at any moment. Luckily, Immortex has a new cure for aging and serious illness – a "mindscan." Immortex scientists will upload his brain configuration into a

computer and "transfer" it into an android body that is designed using his own body as a template. Although imperfect, the android body has its advantages – as the transhumanist FAQ notes, once an individual is uploaded, a backup exists that can be downloaded if one has an accident. And it can be upgraded as new developments emerge. Jake will be immortal.

Sullivan enthusiastically signs numerous legal agreements. He is told that, upon uploading, his possessions will be transferred to the android, who will be the new bearer of his consciousness. Sullivan's original copy, which will die soon anyway, will live out the remainder of his life on "High Eden," an Immortex colony on the moon. Although stripped of his legal identity, the original copy will be comfortable there, socializing with the other originals who are also still confined to biological senescence.

While lying in the scanning tube a few seconds before the scan, Jake reflects:

> I was looking forward to my new existence. Quantity of life didn't matter that much to me – but quality! And to have time – not only years spreading out into the future, but time in each day. Uploads, after all, didn't have to sleep, so not only did we get all those extra years, we got one-third more productive time. The future was at hand. Creating another me. Mindscan.
>
> But then, a few seconds later:
>
> "All right, Mr. Sullivan, you can come out now." It was Dr. Killian's voice, with its Jamaican lilt.
> My heart sank. No . . .
> "Mr. Sullivan? We've finished the scanning. If you'll press the red button . . ."
> It hit me like a ton of bricks, like a tidal wave of blood. No! I should be somewhere else, but I wasn't. . . .
> I reflexively brought up my hands, patting my chest, feeling the softness of it, feeling it raise and fall. Jesus Christ!
>
> I shook my head. "You just scanned my consciousness, making a duplicate of my mind, right?" My voice was sneering. "And since I'm aware of things after you finished the scanning, that means I – this version – isn't that copy. The copy doesn't have to worry about becoming a vegetable anymore. It's free. Finally and at last, it's free of everything that's been hanging over my head for the last twenty-seven years. We've diverged now, and the cured me has started down its path. But this me is still doomed. . . ." (Sawyer 2005: 44–5)

Sawyer's novel is a *reductio ad absurdum* of the patternist conception of the person. For all that patternism says is that as long as person A has the same computational configuration as person B, A and B are the same person. Indeed, Sugiyama, the person selling the mindscan to Jake, had espoused a form of

patternism (Sawyer 2005: 18). Jake's unfortunate experience can be put into the form of a challenge to patternism, which we shall call the "reduplication problem": only one person can really be Jake Sullivan, as Sullivan reluctantly found out. But according to patternism, both creatures are Jake Sullivan – for they share the very same psychological configuration. But, as Jake learned, while the creature created by the mindscan process may be a person, it is not the *very same* person as Jake. It is just another person with an artificial brain and body configured like the original. Hence, having a particular type of pattern cannot be *sufficient* for personal identity. Indeed, the problem is illustrated to epic proportions later in the book when numerous copies of Sullivan are made, all believing they are the original! Ethical and legal problems abound.

A Response to the Reduplication Problem

Perhaps there is a way around this objection. As noted, the reduplication problem suggests that sameness of pattern is not sufficient for sameness of person. However, consider that there seems to be something right about patternism – for as Kurzweil notes, our cells change continually; it is only the organizational pattern that carries on. *Given this, materialism either leaves us with a view of persons in which persons do not persist, or it covertly depends on the idea that we consist in some sort of pattern of organization and is not really a materialist theory at all*. Unless one has a religious conception of the person, and adopts the soul theory, patternism seems inevitable, at least insofar as one believes there is such a thing as a person to begin with. In light of this, perhaps one should react to the reduplication case in the following way: one's pattern is *essential* to one's self despite not being *sufficient for* a complete account of one's identity. Perhaps there is an additional essential property which, together with one's pattern, yields a complete theory of personal identity. But what could the missing ingredient be? Intuitively, it must be a requirement that serves to rule out mindscans and, more generally, any cases in which the mind is "uploaded". For any sort of uploading case will give rise to a reduplication problem, for uploaded minds can in principle be downloaded again and again.

Now, think about your own existence in space and time. When you go out to get the mail, you move from one spatial location to another, tracing a path in space. A spacetime diagram can help us visualize the path one takes throughout one's life. Collapsing the three spatial dimensions into one (the vertical axis) and taking the horizontal axis to signify time, consider the following typical trajectory (Figure 19.1). Notice that the figure carved out looks like a worm; you, like all physical objects, carve out a sort of "spacetime worm" over the course of your existence.

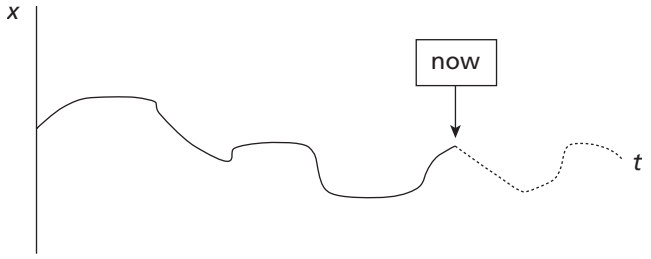

Figure 19.1

This, at least, is the kind of path that "normals" – those who are neither post-humans nor superintelligences – carve out. But now consider what happened during the mindscan. Again, according to patternism, there would be two of the very same person. The copy's spacetime diagram would look like the following:

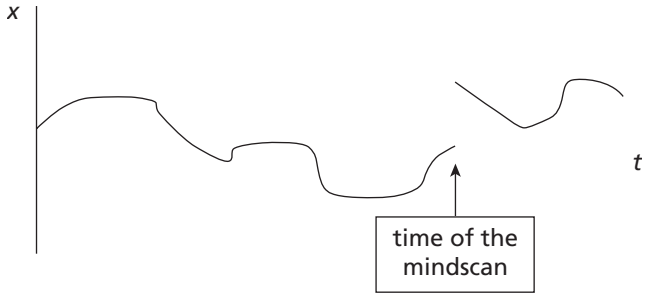

Figure 19.2

This is bizarre. It appears that Jake Sullivan exists for 42 years, has a scan, and then somehow instantaneously moves to a different location in space and lives out the rest of his life! This is radically unlike normal survival. This alerts us that something is wrong with pure patternism: it lacks a requirement for spatiotemporal continuity.

This additional requirement would seem to solve the reduplication problem. For consider the day of the mindscan. Jake went into the laboratory and had a scan; then he left the laboratory and went directly into a spaceship and flew to Mars. It is this man – the one who traces a continuous trajectory through space and time – who is in fact the true Jake Sullivan.

This response to the reduplication problem only goes so far, however. For consider Sugiyama, who, when selling his mindscan product, ventured a patternist pitch. If Sugiyama had espoused patternism together with a spatiotemporal continuity clause, few would have signed up for the scan! For that extra ingredient would rule out a mindscan, or any kind of uploading for that matter, as a form of survival. Only those wishing to have a mere replacement for

themselves would sign up. There is a general lesson here for the transhumanist: if one opts for patternism, enhancements like uploading to avoid death or to facilitate further enhancements easier to carry out, are not really "enhancements" but forms of suicide. *The upshot: the transhumanist should sober up and not offer such procedures as enhancements*. When it comes to enhancement, there are intrinsic limits to what technology can deliver. (Ironically, the proponent of the soul theory is in better shape here. For perhaps the soul does teleport. Who knows?)

Let me sum up the dialectical situation thus far: we have just discarded the original form of patternism as false. If the transhumanist would like to uphold patternism, then she should add the spatiotemporal continuity clause. And importantly, she will need to modify her views on what sorts of enhancements are compatible with survival. Let us call this new position "modified patternism." As we shall now see, although modified patternism is a clear improvement, it requires far more spelling out in at least the following two dimensions.

Two Issues that Modified Patternism Needs to Address

(1) Consider: if you are your pattern, what if your pattern shifts? Do you die? In order for the transhumanist to justify the sort of enhancements needed to become a posthuman or a superintelligent being, she will need to say precisely what a "pattern" is, and when enhancements do and do not constitute a continuation of the pattern. The extreme cases seem clear – for instance, as discussed, mindscans are ruled out by the spatiotemporal continuity clause. And further, because patternism is a psychological continuity view, the patternist will want to say that a memory erasure process that erased one's childhood is an unacceptable alteration of one's pattern, removing too many of one's memories. On the other hand, mere everyday cellular maintenance by nanobots to overcome the slow effects of aging would, according to proponents of this view, not affect the identity of the person.[12] But the middle range cases are unclear. Maybe deleting a few bad chess-playing habits is kosher, but what about erasing all memory of some personal relationship, as in the film *Eternal Sunshine of the Spotless Mind*? The path to superintelligence may very well be a path through middle range enhancements. So again, what is needed is a clear conception of what a pattern is, and what changes in pattern are acceptable and why. Without a firm handle on this issue, the transhumanist developmental trajectory is perhaps the technophile's alluring path to suicide.

This problem looks hard to solve in a way that is compatible with preserving the very idea that we can be identical over time to some previous or future self. For determining a boundary point seems a rather arbitrary exercise in which once a boundary is selected, an example is provided suggesting the boundary

should be pushed outward, ad nauseum. On the other hand, there is something insightful about the view that over time one gradually becomes less and less like one's earlier self. But appreciate this point too long and it may lead to a dark place: for if one finds patternism compelling to begin with, how is it that one truly persists over time, from the point of infancy until maturity, during which time there are often major changes in one's memories, personality, and so on? Indeed, even a series of gradual changes cumulatively amounts to an individual, B, who is greatly altered from her childhood self, A. Why is there really a relation of identity that holds between A and B, instead of an ancestral relation: *A's being the ancestor of B*? Our second issue relates to the issue of gradual, but cumulatively significant, change as well.

(2) Suppose that it is 2050, and people are getting gradual neural regeneration procedures as they sleep. During their nightly slumbers, nanobots slowly import nanoscale materials that are computationally identical to the original materials. The nanobots then gradually remove the old materials, setting them in a small container beside the person's bed. By itself, this process is unproblematic for modified patternism. But now suppose there is an optional upgrade to the regeneration service for those who would like to make a backup copy of their brains. If one opts for this procedure, then, during the nightly process, the nanobots take the replaced materials out of the dish and place them inside a cryogenically frozen biological brain. At the end of the slow process the materials in the frozen brain have been entirely replaced by the person's original neurons. Now, suppose you choose to undergo this procedure alongside your nightly regeneration. Over time, this second brain comes to be composed of the very same material as your brain originally was, configured in precisely the same manner. Which one is you? The original brain, which now has entirely different neurons, or the one with all your original neurons?[13]

The modified patternist has this to say about the neural regeneration case: you are the creature with the brain with entirely different matter, as this creature traces a continuous path through spacetime. But now, things go awry: why is spatiotemporal continuity supposed to outweigh other factors, like being composed of the original material substrate? Here, to be blunt, my intuitions crap out. We'd like to find a solid justification for selecting one option above the other. Until the transhumanist provides a solid justification for her position, it is best to regard forms of enhancement that involve the rapid or even gradual replacement of parts of one's brain as being risky.

Conclusion

I hope all this has convinced you that if the transhumanist maintains patternism there are some serious issues that require working out. Indeed, as *The Transhumanist*

Frequently Asked Questions indicates, the development of radical enhancements, such as brain–machine interfaces, cryogenic freezing for life extension, and uploading to avoid death or simply to facilitate enhancement, are key enhancements invoked by the transhumanist view of the development of the person. Now all of these enhancements sound strangely like the thought experiments philosophers have used for years as problem cases for various theories of the nature of persons, so it is not surprising that deep problems emerge. Herein, I've argued that the Mindscan example suggests that one should not upload and that the patternist needs to modify her theory to rule out such cases. Even with this modification in hand, however, transhumanism still requires a detailed account of what constitutes a break in a pattern versus a mere continuation of it. Without progress on this issue, it will not be clear if medium range enhancements, such as erasing childhood memories or adding neural circuitry to make oneself smarter, are safe. Finally, the nanobot case warns against even mild enhancements. Given all this, it is fair to say that the transhumanist currently cannot support her case for enhancement. Indeed, *The Transhumanist Frequently Asked Questions* notes that transhumanists are keenly aware that this issue has been neglected:

> While the concept of a soul is not used much in a naturalistic philosophy such as transhumanism, many transhumanists do take an interest in the related problems concerning personal identity (Parfit 1984) and consciousness (Churchland 1988). These problems are being intensely studied by contemporary analytic philosophers, and although some progress has been made, e.g. in Derek Parfit's work on personal identity, they have still not been resolved to general satisfaction. (Bostrom 2003: section 5.4)

Our discussion also raises some general lessons for all parties involved in the enhancement debate. For when one considers the enhancement debate through the lens of the metaphysics of personhood, new dimensions of the debate are appreciated. The literature on the nature of persons is extraordinarily rich, raising intriguing problems for commonly accepted views of the nature of persons that underlie positions on enhancement. When one defends or rejects a given enhancement, it is important to determine whether one's stance on the enhancement in question is truly supported by, or even compatible with, one's position on the nature of persons. Further, the topic of the nature of persons is of clear relevance to the related topics of human nature and human dignity, issues that are currently key points of controversy in debates over enhancement (see, e.g., Bostrom 2008; Fukuyama 2002).

Perhaps, alternately, you grow weary of all this metaphysics. You may suspect that social conventions concerning what we commonly consider to be persons are all we have because metaphysical theorizing will never conclusively resolve what persons are. However, as unwieldy as metaphysical issues are, it seems

that not all conventions are worthy of acceptance, so one needs a manner of determining which conventions should play an important role in the enhancement debate and which ones should not. And it is hard to accomplish this without getting clear on one's conception of persons. Further, it is difficult to avoid at least implicitly relying on a conception of persons when reflecting on the case for and against enhancement. For what is it that ultimately grounds your decision to enhance or not to enhance, if not that it will somehow improve you? Are you perhaps merely planning for the well-being of your successor?

Notes

1. This piece is expanded and modified from an earlier piece, "Future Minds: Cognitive Enhancement, Transhumanism, and the Nature of Persons", which appeared in the *Penn Bioethics Reader*, Vardit Radvisky and Art Caplan (eds.), in press. Thanks very much to Ted Sider and Michael Huemer for their helpful comments.
2. Julian Huxley apparently coined the term *transhumanism* in 1957, when he wrote that in the near future "the human species will be on the threshold of a new kind of existence, as different from ours as ours is from that of Peking man" (Huxley 1957: 13–17).
3. Bostrom is a philosopher at Oxford University who now directs the transhumanist-oriented Future of Humanity Institute there. In addition to these two documents, there are a number of excellent philosophical and sociological works that articulate key elements of the transhumanist perspective (e.g. Bostrom 2005; Hughes 2004; Kurzweil 1999, 2005). For extensive Web resources on transhumanism, see Nick Bostrom's homepage, Ray Kurzweil's newsgroup (KurzweilAI.net), the Institute for Ethics and Emerging Technologies homepage, and the World Transhumanist Association homepage.
4. It should be noted that transhumanism by no means endorses every sort of enhancement. For example, Nick Bostrom rejects positional enhancements (enhancements primarily employed to increase one's social position) yet argues for enhancements that could allow humans to develop ways of exploring "the larger space of possible modes of being" (2005: 11).
5. There are many nuances to this rough trajectory. For instance, some transhumanists believe that the move from unenhanced human intelligence to superintelligence will be extremely rapid because we are approaching a singularity, a point at which the creation of superhuman intelligence will result in massive changes in a very short period (e.g. 30 years) (Bostrom 1998; Kurzweil 1999, 2005; Vinge 1993). Other transhumanists hold that technological changes will not be so sudden. These discussions often debate the reliability of Moore's Law (Moore 1965). Another key issue is whether a transition to superintelligence will really occur because the upcoming technological developments involve grave risk. The risks of biotechnology and AI concern transhumanists, progressive bioethicists more generally, as well as bioconservatives (Annis 2000; Bostrom 2002; Garreau 2005; Joy 2000).

6. For mainstream anti-enhancement positions on this question see, e.g., Fukuyama (2002), Kass et al. (2003), and Annas (2000).
7. Because our discussion is introductory, I will not delve into different versions of psychological continuity theory. One could, for instance, appeal to (a): the idea that memories are essential to a person. Alternatively, one could adopt (b), one's overall psychological configuration is essential, including one's memories. Herein, I shall work with one version of this latter conception – one that is inspired by cognitive science – although many of the criticisms of this view will apply to (a) and other versions of (b) as well. For some different versions see chapter 27 of John Locke's 1694 *Essay Concerning Human Understanding* (note that this chapter first appears in the second edition; it is also reprinted as "Of Identity and Diversity" in Perry 1975). See also the essays by Anthony Quinton and Paul Grice, both of which are reprinted in Perry (1975).
8. Sociologist James Hughes holds a transhumanist version of the no self view. (See the Institute for Ethics and Emerging Technology's "Cyborg Buddha" project at http://ieet.org/index.php/IEET/cyborgbuddha,) For helpful surveys of these four positions, see Eric Olson's chapter in this volume (Chapter 7), and Conee and Sider (2005).
9. It should be noted that although a number of bioconservatives seem to uphold the soul theory, the soul theory is not, in and of itself, an anti-enhancement position. For why can't one's soul or immaterial mind inhere in the same body even after radical enhancement?
10. For discussion of computational theories, see Block (Chapter 14 in this volume) and Churchland (1996).
11. Readers familiar with philosophy of mind may suggest that the transhumanist could accept one version of materialism, namely, "token materialism." However, I suspect that it is not really a coherent form of materialism. Token materialism holds that every instance of a mental property is identical to some instance of a physical property. But can the instances really be *identical* if the properties themselves belong to different types? The distinct property types are instead coinstantiated by the same particular.
12. Or at least, this is what the patternist would *like* to say. The example in the next paragraph will in fact question whether she can truly say this.
13. This is a science fiction variant of the well-known Ship of Theseus case. It first appears in print in Plutarch (*Vita Thesei*, 22–3).

References

Annas, G. J. (2000). The man on the moon, immortality, and other millennial myths: The prospects and perils of human genetic engineering. *Emory Law Journal* 49 (3): 753–82.

Bostrom, N. (1998). How long before superintelligence? *International Journal of Futures Studies* 2. Available at http://www.nickbostrom.com/superintelligence.html (retrieved Dec. 20, 2008).

Bostrom, N. (2003). *The Transhumanist Frequently Asked Questions*: v 2.1. World Transhumanist Association. Retrieved from http://transhumanism.org/index.php/WTA/faq/.

Bostrom, N. (2005). History of Transhumanist Thought. *Journal of Evolution and Technology*, 14 (1).

Bostrom, N. (2008). Dignity and enhancement. In The President's Council on Bioethics, *Human Dignity and Bioethics: Essays Commissioned by the President's Council on Bioethics* (Washington, DC: U.S. Government Printing Office).

Conee, E. and Sider, T. (2005). *Riddles of Existence: A Guided Tour of Metaphysics*. Oxford: Oxford University Press.

Churchland, P. (1988). *Matter and consciousness*. Cambridge, MA: MIT Press.

Churchland, P. (1996). *Engine of reason, seat of the soul*. Cambridge, MA: MIT Press.

Fukuyama, F. (2002). *Our posthuman future: Consequences of the biotechnology revolution*. New York Farrar, Straus and Giroux.

Garreau, J. (2005). *Radical evolution: The promise and peril of enhancing our minds, our bodies – and what it means to be human*. New York: Doubleday.

Hughes, J. (2004), *Citizen cyborg: Why democratic societies must respond to the redesigned human of the future*. Cambridge, MA: Westview Press.

Huxley, J. (1957). *New bottles for new wine*. London: Chatto & Windus.

Joy, B. (2000). Why the future doesn't need us. *Wired*, 8: 238–46.

Kurzweil, R. (1999). *The age of spiritual machines: When computers exceed human intelligence*. New York: Viking.

Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. New York: Viking.

Moore, G. (1965). Cramming more components into integrated circuits. *Electronics*, 38 (8): 11–17. Retrieved from ftp://download.intel.com/research/silicon/moorespaper.pdf.

Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.

Perry, J. (1975). *Personal identity*, Berkeley: University of California Press.

Roco, M. C., and Bainbridge W. S. (eds.) (2002). *Converging technologies for improved human performance: Nanotechnology, biotechnology, information technology and cognitive science*. Arlington, VA: National Science Foundation/Department of Commerce.

Sawyer, R. (2005). *Mindscan*. New York: Tor.

Vinge, V. (1993). The coming technological singularity: How to survive in the post-human era. *NASA Technical Reports*, Lewis Research Center, Vision 21: Interdisciplinary Science and Engineering in the Era of Cyberspace, pp. 11–21.

World Transhumanist Association (1998). *Transhumanist Declaration*. Retrieved from http://www.transhumanism.org/index.php/WTA/declaration/.