



## Vulnerable Cyborgs: Learning to Live with our Dragons

Mark Coeckelbergh  
Department of Philosophy  
University of Twente  
[m.coeckelbergh@utwente.nl](mailto:m.coeckelbergh@utwente.nl)

*Journal of Evolution and Technology* - Vol. 22 Issue 1 – November 2011 - pgs 1-9

### Abstract

Transhumanist visions appear to aim at invulnerability. We are invited to fight the dragon of death and disease, to shed our old, human bodies, and to live on as invulnerable minds or cyborgs. This paper argues that even if we managed to enhance humans in one of these ways, we would remain highly vulnerable entities given the fundamentally relational and dependent nature of posthuman existence. After discussing the need for minds to be embodied, the issue of disease and death in the infosphere, and problems of psychological, social and axiological vulnerability, I conclude that transhumanist human enhancement would not erase our current vulnerabilities, but instead transform them. Although the struggle against vulnerability is typically human and would probably continue to mark posthumans, we had better recognize that we can never win that fight and that the many dragons that threaten us are part of us. As vulnerable humans and posthumans, we are at once the hero and the dragon.

### Introduction

Transhumanists have articulated visions that seem to aim at invulnerability and immortality. Consider the writings of two well-known proponents of human enhancement: Nick Bostrom and Ray Kurzweil. Bostrom has written a tale about a dragon that terrorizes a kingdom and people who submit to the dragon rather than fighting it. According to Bostrom, the “moral” of the story is that we should fight the dragon, that is, extend the (healthy) human life span and not accept aging as a fact of life (Bostrom 2005, 277). And in *The Singularity is Near* (2005) Kurzweil has suggested that following the acceleration of information technology, we will become cyborgs, upload ourselves, have nanobots in our bloodstream, and enjoy nonbiological experience. Although not all transhumanist authors explicitly state it, these ideas seem to aim toward invulnerability and immortality: by means of human enhancement technologies, we can transcend our present limited existence and become strong, invulnerable cyborgs or immortal minds living in an eternal, virtual world.

Given these aims, the mythical and religious language used in transhumanist fables and visions is entirely appropriate. For instance, many ideas about human enhancement appear to fit a Gnostic way of thinking: we should leave the material world, transcend our earthly bodies, and resurrect into the eternal life of non-material existence. And just as (other) religious ideas, transhumanist visions have created much ethical controversy. Is it right for humans to enhance themselves?

However, in this paper, I will ask neither the ethical-normative question (Should we develop human enhancement techniques and should we aim for invulnerability?) nor the hermeneutical question (How can we best interpret and understand transhumanism in the light of cultural, religious, and scientific history?). Instead, I ask the question: *If and to the extent that* transhumanism aims at invulnerability, can it – in principle – reach that aim? The following discussion offers some obvious and some much less obvious reasons why posthumans would remain vulnerable, and why human vulnerability would be transformed rather than diminished or eliminated.

Note that this is a different problem from asking about the *value* of vulnerability, that is, the value of human vulnerability. For instance, based on Nussbaum's work one could set up a post-Stoic defence of human vulnerability: the same external realities that make us vulnerable can also be a source of human value, and valuing makes us dependent (Nussbaum 1986) (I will return to this point below). More generally, in response to transhumanism one could offer objections based on what one thinks is valuable in current human existence. The discussion below is relevant to such a project. However, to focus only on a defense or rejection of what is valuable in humans would leave out of sight the relation between (in)vulnerability and *posthuman* possibilities. It would lead us back to the ethical-normative questions (Is human enhancement morally acceptable? Is vulnerability something to be valued? Is the transhumanist project acceptable or desirable?), which is not what I want to do in this paper. Moreover, ethical arguments that present the problem as if we have a choice between "natural" humanity and "artificial" posthumanity are based on essentialist assumptions that make a sharp distinction between "what we are" (the natural) and technology (the artificial), whereas this distinction is at least questionable. Perhaps there is no fixed human nature apart from technology, perhaps we are "artificial by nature" (Plessner 1975). If this is so, then the problem is not whether or not we want to transcend the human but how we want to shape that posthuman existence. Should we aim at invulnerability and if so, can we? As indicated before, here I limit the discussion to the "can" question. In my conclusion, however, I will return to the issue of how to formulate the ethical question concerning human enhancement.

### **Sources of vulnerability: Some posthuman dragons**

Let me list and discuss some sources of posthuman vulnerability. The categories are not meant to be mutually exclusive or to mark strongly distinctive domains; by using them I highlight different aspects of posthuman vulnerability in order to show that in so far as transhumanism aims at invulnerability, it must necessarily fail.

Note that the same arguments can be made for current humans "enhanced" by contemporary technology, for instance, as compared to early humans and other biological ancestors. I will often refer to human vulnerability and make comparisons between humans and posthumans, but the focus will be on posthumans.

### *Physical vulnerability*

As transhumanists will agree, no form of human enhancement ensures full protection against physical threats: posthumans can always be harmed by other posthumans or by external forces that are not within posthuman control. Of course human enhancement would offer new protections against specific threats. There would be new immunities. Not only could human enhancement make us immune to current viruses; it could also offer other “immunities,” broadly understood. For instance, we might consider changes that enable us to deal better with threats to our emotional well-being. (I will return to psychological vulnerability below.) However, the project of total vulnerability or even overall reduction of vulnerability is bound to fail. If we consider the history of medical technology, we observe that for every disease new technology helps to prevent or cure, there is at least one new disease that escapes our techno-scientific control. We can win one battle, but we can never win the war. There will be always new diseases, new viruses, and, more generally, new threats to physical vulnerability. Consider also natural disasters caused by floods, earthquakes, volcanic eruptions, and so on.

Moreover, the very means to fight those threats sometimes create new threats themselves. This can happen within the same domain, as is the case with antibiotics that lead to the development of more resistant bacteria, or in another domain, as is the case with new security measures in airports, which are meant as protections against physical harm by terrorism but might pose new (health?) risks. Paradoxically, technologies that are meant to reduce vulnerability often create new ones. This is also true for posthuman technologies. For example, posthumans would also be vulnerable to at least some of the risks Bostrom calls “existential risks” (Bostrom 2002), which could wipe out posthumankind. Nanotechnology or nuclear technology could be misused, a superintelligence could take over and annihilate humankind, or technology could cause (further) resource depletion and ecological destruction. Military technologies are meant to protect us but they can become a threat, making us vulnerable in a new way. We wanted to master nature in order to become less dependent on it, but now we risk destroying the ecology that sustains us. And of course there are many physical threats we cannot foresee – not even in the near future. Posthumans will remain vulnerable to at least some existing physical threats, but they will also face new risks and create new vulnerabilities.

### *Material and immaterial vulnerability*

Physical vulnerability is not limited to threats to the human body. We have always extended that body with technology. This process of “cyborgization” is likely to increase in posthumans, who would extend themselves with information technology and other technology to a much higher degree than contemporary humans. But whether or not we want to use the term “cyborg” for this, as Haraway and Clark do (Haraway 1991; Clark 2003), taking seriously the idea that we are extending our body and mind with technology implies that we should broaden the concept of vulnerability as well. For instance, today we already extend our bodies and minds by means of information technology. But that technology is also vulnerable to threats of various kinds. Consider computer viruses. Here the story is similar to the story of biological viruses: there are ongoing cycles of threats, counter-measures, and new threats. We can also consider physical damage to computers, although that is much less common. In any case, if we extend ourselves with software and hardware, this creates additional vulnerabilities. We must cope with “software” vulnerability and “hardware” vulnerability. If humans and posthumans live in an “infosphere” (see for example Floridi 2002), this is not a sphere of immunity. Perhaps our vulnerability becomes less material, but we cannot escape it.

For instance, a virtual body in a virtual world may well be shielded from biological viruses, but it is vulnerable to at least three kinds of threats. First, there are threats within the virtual world itself (consider for instance virtual rape), which constitutes virtual vulnerability. Second, the software programme that provides a platform for the virtual world might be damaged, for example by means of a cyber attack. This can lead to the “death” of the virtual character or entity. Third, all these processes depend on (material) hardware. The world wide web and its wired and wireless communications rest on material infrastructures without which the web would be impossible. Therefore, if posthumans uploaded themselves into an infosphere and dispensed with their biological bodies, they would not gain invulnerability and immortality but merely transform their vulnerability. Their “ghostly” existence would be highly vulnerable given these and other immaterial and material threats.

### *Bodily vulnerability*

Fantasies about immaterial and invulnerable existence in the infosphere do not generally try to cancel out the body. This is because we *need* one. Minds need bodies. This is in line with contemporary research in cognitive science, which argues that “embodiment” is necessary since minds can develop and function only in interaction with their environment (Lakoff and Johnson 1999 and others). This direction of thought is also taken in contemporary robotics, for example when it recognizes that manipulation plays an important role in the development of cognition (Sandini et al. 2004). In his famous 1988 book on “mind children” Moravec argued that true AI can be achieved only if machines have a body (Moravec 1988). This is also acknowledged by Kurzweil, albeit for a different, more superficial reason which seems to assume that posthumans must resemble humans:

If we are truly capturing a particular person’s mental processes, then the reinstated mind will need a body, since so much of our thinking is directed toward physical needs and desires. [...] The human body version 2.0 will include virtual bodies in completely realistic virtual environments, nanotechnology-based physical bodies, and more. (Kurzweil 2005, 199)

Thus, uploading and nano-based cyborgization would not dispense with the body but transform it into a virtual body or a nano-body. This would create vulnerabilities that sometimes resemble the vulnerabilities we know today (for instance virtual violence) but also new vulnerabilities. For instance, no one knows what kinds of risks would emerge if we had nano robots in our blood stream. Our bodies would be transformed in ways that are hard to imagine, and so would our vulnerability.

### *Metaphysical vulnerability*

According to an influential metaphysical doctrine, bodies are organizations of matter, in particular organizations of elementary particles. The particular combinations of matter are always temporary since they are vulnerable to disintegration. The Greek philosopher Democritus, known as the founder of atomism, claimed that whereas atoms are eternal, the objects composed of them are not. Worlds come and disappear again. And while contemporary physics and metaphysics are no longer atomist in the common sense of the word (“atoms” turned out not to be the smallest particle), physics is still after elementary particles and the natural sciences embrace an atomist metaphysics concerning the relation between systems (or organisms) and their elements. Even the “infosphere” (Floridi 2002) has

its “information objects,” which might be interpreted as compositions of “elementary particles”: bits.

With this atomism comes that atomist view of death: there is always the possibility of disintegration; neither physical-material objects nor information objects exist forever. Information can disintegrate and the material conditions for information are vulnerable to disintegration as well. Thus, at a fundamental level everything is vulnerable to disintegration, understood by atomism as a re-organization of elementary particles. This “metaphysical” vulnerability is unavoidable for posthumans, whatever the status of their elementary particles and the organs and systems constituted by these particles (biological or not). According to their own metaphysics, the cyborgs and inforgs that transhumanists and their supporters wish to create would be only temporal orders that have only temporary stability – if any.

Note, however, that recently both Floridi and contemporary physics seem to move toward a more ecological, holistic metaphysics, which suggests a different definition of death. In information ecologies, perhaps death means the absence of relations, disconnection. Or it means: deletion, understood ecologically and holistically as the removal out of the whole. But in the light of this metaphysics, too, there seems no reason why posthumans would be able to escape death in this sense. Whether they are seen as composed of elementary particles or as relational nodes in a network-ecology, they remain vulnerable and “mortal,” however virtual they might have become.

#### *Existential and psychological vulnerabilities*

Vulnerability has its source not only in material-ontological reality, but also in existential experience, psychology, and perception. We are not only directly vulnerable as bodily, material, and (meta)physical entities; as humans we can also *know* and *experience* those vulnerabilities. This gives rise to what we may call “indirect” or “second-order” vulnerabilities. For instance, we can become aware of the possibility of disintegration, the possibility of death. We can also become aware of less threatening risks, such as disease. There are many first-order vulnerabilities. Awareness of them renders us extra vulnerable as opposed to beings who lack such an ability to take distance from ourselves. From an existential-phenomenological point of view (which has its roots in work by Heidegger and others), but also from the point of view of common sense psychology, we must extend the meaning of vulnerability to the sufferings of the mind. Vulnerability awareness itself constitutes a higher-order vulnerability that is typical of humans. In posthumans, we could only erase this vulnerability if we were prepared to abandon the particular higher form of consciousness that we “enjoy.” No transhumanist would seriously consider that solution to the problem. Therefore, if posthumans were to have a higher form of consciousness not too dissimilar to ours, then they would have to cope with second-order vulnerabilities as well as first-order ones.

#### *Social and emotional vulnerability*

We do not live in isolation: we are social beings who depend on each other for fulfilling our physical, emotional, and other needs, and this makes us vulnerable in many ways. For example, we tend to form relationships, groups and communities and along with the many advantages that this offers, it also produces plenty of possibilities for suffering and violence. If I depend on you socially and emotionally, then I am vulnerable to what you say or do. Unless posthumans were to live in complete isolation without any possibility of inter-

posthuman communication, they would be as vulnerable as we are to the sufferings created by the social life, although the precise relation between their social life and their emotional make-up might differ.

An interesting vision to study in this respect is the one suggested by Houellebecq in his novel *The Possibility of an Island* (2005). In the story, genetic and other enhancement interventions abolish society as we know it. Instead, the novel projects posthumans spending their lives in isolation, as hermits living in “compounds” that are fenced off from the harsh natural world and from “degenerated” humans who revert back to “primitive” and violent forms of group life. However, the posthumans can still communicate and relate to their “ancestors” through reading and writing. (If this is a “transhumanist” vision at all, it is still humanist in the sense that it preserves the reading/writing of humanism understood as a *writing* movement, a movement that centers on the technology of writing. Moreover, “ancestors” takes on a different meaning since their “descendants” are clones.) Thus, the tension we modern humans know between trying to reach immunity and experiencing being caught up in social-relational dependency, remains in place. Only fiercely anti-humanist enhancement would abolish social relations entirely.

Of course even in a non-isolationist vision, posthumans might be changed in such a way that they would have a different emotional life. For example, in Houellebecq’s novel the posthumans have a reduced capacity to feel sad, but at the cost of a reduced capacity to desire and to feel joy. More generally, the lesson seems to be: emotional enhancement comes at a high price. Are we prepared to pay it? Even if we succeed in diminishing this kind of vulnerability, we might lose something that is of value to us. This brings me to the next kind of vulnerability.

### *Ethical-axiological vulnerability*

Humans are not just witnesses and interpreters of physical and social processes. They also evaluate the processes and engage with them. But the very activity of *valuing* renders us vulnerable. We value not only people and our relationships with them; we are also attached to many other things in life. Caring makes us vulnerable (Nussbaum 1986). We develop ties out of our engagement with humans, animals, objects, buildings, landscapes, and many other things. This renders us vulnerable since it makes us dependent on (what we experience as) “external” things. We sometimes get emotional about things since we care and since we value. We *suffer* since we depend on external things. Valuing is a source of joy but also of harm. The Stoics knew this and followed a particular strategy of immunity: they tried to disarm the emotions and the vulnerability by not caring about the externalities, that is, by trying to cut the ties, the dependencies.

Posthumans could be cognitively equipped to follow this strategy, for instance by means of emotional enhancement that allows more self-control and prevents them forming too strong ties to things. If we really wanted to become invulnerable in this respect, we should create posthumans who no longer care at all about external things – including other posthumans. That would be “posthumans” who no longer have the ability to care and to value. They would “connect” to others and to things, but they would not really engage with them, since that would render them vulnerable. They would be perfectly rational Stoics, perhaps, but it would be odd to call them “posthumans” at all since the term “human” would lose its meaning. It is even doubtful if this extreme form of Stoicism would be *possible* for any entity that possesses the capacity of valuing and that *engages* with the world. Again, transhumanists could render

this possible only if they were prepared to give up axiological and emotional ways of engaging with the world. If they wanted to avoid this consequence, they could propose more modest forms of “fine-tuning” to our existing cognitive make-up, without compromising the human capacity to care and value. However, this implies that posthumans would retain a large degree of their ethical-axiological vulnerability.

### *Relational vulnerability*

In sum, because our imagined posthumans remain relational beings, operating in a web of dependencies without which they could not exist, they remain vulnerable in various ways. They are dependent on their physical environment, on their bodies, on the technological and biological systems that embody and extend their minds, on other posthumans and on the people and things they value. The only way to make an entity invulnerable, it turns out, would be to create one that exists in absolute isolation and is absolutely independent of anything else. Such a being seems inconceivable – or would be a particularly strange kind of god. (It would have to be a “philosopher’s” god that could hardly stir any religious feelings. Moreover, the god would not even be a “first mover,” let alone a creator, since that would imply a relation to our world. It is also hard to see how we would be aware of its existence or be able to form an idea about it, given the absence of *any* relation between us and the god.) Of course we could – if ethically acceptable at all – create posthumans that are less vulnerable in some particular areas, as long as we keep in mind that there are other sources of vulnerability, that new sources of vulnerability will emerge, and that our measure to decrease vulnerability in one area may increase it in another area.

If transhumanists accept the results of this discussion, they should carefully reflect on, and redefine, the aims of human enhancement and avoid confusion about how these aims relate to vulnerability. If the aim is invulnerability, then I have offered some reasons why this aim is problematic. If their project has nothing to do with trying to reach invulnerability, then why should we transcend the human? Of course one could formulate no “ultimate” goals and choose less ambitious goals, such as more health and less suffering. For instance, one could use a utilitarian argument and say that we should avoid overall suffering and pain. Harris seems to have taken these routes (Harris 2007). And Bostrom frequently mentions “life extension” as a goal rather than “invulnerability” or “immortality.” But even in these “weakened” or at least more modest forms, the transhumanist project can be interpreted as a particularly hostile response to (human) vulnerability that probably has no parallel in human history. Making people less vulnerable remains an important goal of transhumanists, in spite of their likely acknowledgment that *absolute* vulnerability is impossible. This means that the limitations discussed above remain highly relevant to their project and cannot be dismissed as obvious or off the mark. They help to answer the question: Is invulnerability one of the aims of transhumanist human enhancement, and if so, *how* invulnerable do we want to become and what *kind* of invulnerabilities do we want to achieve?

### **Conclusion: Heels and dragons**

There is a Greek myth that tells us about Achilles who was made invulnerable in his youth – invulnerable except his heel. Tragically, he is said to have died from a heel wound caused by an arrow shot at him. In this paper, I have given several reasons why posthumans would not be unlike Achilles in this respect and why we had better take seriously the ancient Greek sparkle of wisdom when reflecting on human enhancement. In so far as posthuman heroes and their creators might try to transcend vulnerable existence, they would be bound to fail because

there would be many heels, and as we create new technology new heels are created. In facing these tragic cycles of trial and failure, posthumans would be remarkably similar to their human ancestors, who also struggle against their vulnerable condition and have no choice but to live with that condition.

Furthermore, this paper suggests that if we can and must make an ethical choice at all, then it is not a choice between vulnerable humans and invulnerable posthumans, or even between vulnerability and invulnerability, but a choice between different *forms* of humanity and vulnerability. If implemented, human enhancement technologies such as mind uploading will not cancel vulnerability but transform it. As far as ethics is concerned, then, what we need to ask is which new forms of the human we want and how (in)vulnerable we wish to be. But this inquiry is possible only if we first fine-tune our ideas of what is possible in terms of enhancement and (in)vulnerability. To do this requires stretching our moral and technological imaginations.

Moreover, if I'm right about the different forms of posthuman vulnerability as discussed above, then we must dispense with the dragon metaphor used by Bostrom: vulnerability is not a matter of "external" dangers that threaten or tyrannize us, but that have nothing to do with what we are; instead, it is bound up with our relational, technological and transient kind of being – human or posthuman. If there are dragons, they are part of us. It is our tragic condition that as relational entities we are at once the heel and the arrow, the hero and the dragon.

Finally, perhaps it is a consolation for both humans and posthumans that, as Nussbaum suggested, vulnerability is not only a source of suffering but also of joy and value. If flourishing or meaning is what we seek rather than invulnerability, then it seems that now and in the posthuman future we can only find these goods *in* the very dependencies that may sometimes hurt or even destroy us.

## References

Bostrom, N. 2002. Existential risks: Analyzing human extinction scenarios and related hazards. *Journal of Evolution and Technology*, 9, <http://www.jetpress.org.volume9/risks.html>

Bostrom, N. 2005. The fable of the dragon tyrant. *Journal of Medical Ethics* 31(5): 273-77.

Clark, A. 2003. *Natural-born cyborgs: Minds, technologies and the future of human intelligence*. New York: Oxford University Press.

Floridi, L. 2002. On the intrinsic value of information objects and the infosphere. *Ethics and Information Technology* 4(4): 287-304.

Harris, J. 2007. *Enhancing evolution: The ethical case for making better people*. Princeton, NJ: Princeton University Press.

Haraway, D. 1991. A cyborg manifesto: science, technology, and socialist-feminism in the late twentieth century. In *Simians, cyborgs and women: The reinvention of nature*, 149-181. New York: Routledge.

Kurzweil, R. 2005. *The singularity is near: when humans transcend biology*. New York:

Penguin Group.

Lakoff, G. and M. Johnson. 1999. *Philosophy in the flesh: the embodied mind and its challenge to Western thought*. New York: Basic Books.

Moravec, H. 1988. *Mind children: The future of robot and human intelligence*. Cambridge, MA: Harvard University Press

Nussbaum, M.C. 1986. *The fragility of goodness: Luck and ethics in Greek tragedy and philosophy*. Cambridge, MA: Cambridge University Press.

Plessner, H. 1975. *Die Stufen des Organischen und der Mensch: Einleitung in die philosophische Anthropologie*. Berlin: Walter de Gruyter. (Orig. Pub. 1928.)

Sandini, G., G. Metta, and D. Vernon. 2004. RobotCub: An open framework for research in embodied cognition. In *IEEE-RAS/RSJ International Conference on Humanoid Robots (Humanoids 2004)*. November 10-12, 2004 Santa Monica, Los Angeles, CA, USA.