Reframing the Debate on Human Enhancement

EDITED BY
Fiorella Battaglia and Antonio Carnevale

TABLE OF CONTENTS

COMMENTARIES

Epistemological and Moral Problems
with Human Enhancement*

Fiorella Battaglia [†]
fiorella.battaglia@lrz.uni-muenchen.de

Antonio Carnevale [‡]
a.carnevale@sssup.it

We are going to give neither an intensional definition of the concept of human enhancement nor are we going to attempt to build a collection that contains all the possible interventions (extensional). Our analysis draws on an overall meaning of human enhancement as scientific and technological progress that expands the possibilities of human action and reduces its dependence on natural or cultural predetermined constraints, allowing the human condition to be changed via science and technology. Human enhancement thus refers to extended cognitive skills, extended sensory capacities, a significant increase in life expectancy, mood modulation as well as new capabilities that might be provided to healthy individuals.

Rather than dealing with the definitional issue, we aim to sketch out the epistemological and moral underpinnings of the debate on human enhancement in order to provide a demarcation of issues regarding human enhancement. Ultimately, the aim of this issue is to go beyond sterile disputes between supporters and detractors of human enhancement. There are two main difficulties with the current debate. For one thing, it is coined by a discussion of the ambiguous notion of the "natural". The second problem is that the debate centres around naïve attempts to speculate about technology-

[†] Ludwig-Maximilians-Universität, Munich, Germany.
[‡] Scuola Superiore Sant'Anna, Pisa, Italy.

based visions of the future. We advocate a reoriented debate that can complement and inform ongoing work in science and societal debate.

In the remainder, we want to address the following six points that are important for the debate on human enhancement:

1. It is emphasised that we should move past the contemporary discussion on human enhancement and go beyond sterile disputes between its supporters and detractors.

2. It is necessary to interpret the term "human enhancement" in a very wide sense in order to include not just the interventions themselves, but the social and cultural dimension as well.

3. Technologies are not just tools that humans use in order to interact with and experience the surrounding world. They also are means of mediation that shape their world and themselves.

4. The normative – as a way of relating to a standard – is presupposed by our lifeworld and cannot be understood from a purely naturalistic stance.

5. Ethics and law are two forms of the normative. Their relation is analysed in the specific case of human enhancement.

6. Human enhancement affects society at large. Thus it requires public debate and stands in need of regulation.

A discussion of aforementioned points makes it possible to reframe the debate on human enhancement and to achieve a new level of discussion.

As human enhancement roughly refers to numerous interventions that can change in significance depending on the intention with which they are undertaken, we do not want to start with an intensional definition. For example, an intervention to increase the height in an adolescent with normal growth-hormone (GH) secretion whose parents are short is categorized as an enhancement. Instead, the same measure performed on a patient with documented GH deficiency that would result in the same height would be categorized as a therapeutic intervention, thus providing a counterexample to an obvious attempt at an essentialist definition.

The problem with the extensional definition is the heterogeneity of possible interventions, which are based on knowledge from various disciplines. The

scientific knowledge required for the various interventions for the purpose of human enhancement is plentiful, ranging from neuroscience to biology, from engineering to nanotechnology, etc.

The problem therefore seems to move from the definition to the issue of categorization of the interventions of human enhancement. The question thus becomes: how do we categorize human enhancement? How do we decide that a certain intervention is an enhancement? How do we establish that interventions as diverse as a prosthesis, a drug and a technological support as an implant may all be labelled as enhancements? In recent discussion, different scholars have taken different stances on the issue of categorization: some think that our categorical framework should be based on function. Others believe that it should instead be based on the intention. Yet others are convinced that it should be based on a combination of structural properties and function. Thus the issue of definition calls for a consideration of the issue of categorization. In fact, if we think in terms of categorization issues – that is, in terms of comprehending the functional character of human enhancement – then we can improve the quality of debate vis-à-vis the standard discussion of human enhancement that we find in mainstream philosophy, which is focused on definitional issues. On the one hand, it will be much harder then to take an essentialist point of view with respect to the definition of human enhancement. On the other hand, it will also be possible to overcome the difficulties entailed in the distinction between therapy and enhancement that rests on a kind of "between Scylla and Charybdis". Both of which bring problems. The essays included in the present issue integrate concrete, empirical examples and try to move forward the discussion on human enhancement based on these examples. By presenting new empirical findings on the topic, these essays examine how these can lead to philosophical problems.

1. Moving beyond the Pro-Enhancement / Anti-Enhancement Frame

Human enhancement is one of those topics that is likely to polarise. In other words, the first immediate reaction is an emotional (or ideological) reaction between those who are for human enhancement and those who are against human enhancement.

These preferences are supported in various ways by referring to topics such as human dignity or hubris that play a role in the argument. These arguments, however, do not always appear to be in agreement with other beliefs that are also shared by those who support them. There are then two problems: the problem of consistency and the need to clarify conceptually the novelty entailed in emergent technologies. A top-down

approach, that is characteristic of bioethics, has been adopted as a solution. But even this approach has soon been proved to be insufficient. In brief, the shortcomings of a bioethics approach consist in the mere application of established principles a priori that are recognized as valid in all circumstances without regard to the specificity and situatedness of each intervention. Overlooked in this approach are several factors that contribute to determining the specificity of the context. Below, we give a brief but systematic exposition of elements that have to be taken in account in analysing interventions that aim at improving human traits or capacities according to Straub (Straub et al. 2012): 1. the subject; 2. the object; 3. the regard in which the object is enhanced (behaviour, body, personal traits, etc.); 4. the knowledge required for the specific enhancement; 5. the criterion for successful enhancements; 6. the beneficiary of the enhancement; 7. the procedure that is used to bring the about the enhancement; 8. the means that are used; 9. the intended and unintended consequences.

A further difficulty – highlighted by approaches that are critical of the bias inherent in the simple risk assessment and instead promote a multidisciplinary approach called ELSA (Ethical, Legal, and Social Approach) – is that very often we do not have to deal with the concrete improving intervention. We have to deal, rather, with the objectified visions developed in trials and laboratory prototypes. Since the situation is strongly marked by a lack of knowledge, it is difficult to use the criteria of risk assessment.

The analysis of these briefly mentioned points along with the issue of categorization and definition of improving interventions can help to develop a more balanced approach, that goes beyond ideological and emotional responses to the debate on human enhancement.

2. Humanities, Science and Technology

In our intention, reframing the debate on human enhancement also means to address the issue known as the "Two Cultures". The Two Cultures is the title of the first part of an influential Lecture by the British scientist and novelist C. P. Snow. In 1959, Snow pointed to the growing gap between the 'two cultures', between the truth claims of hard science, on the one hand, and the truth claims of the humanities that more closely reflect the human lifeform on the other. More generally speaking, we can say that the term "human enhancement" is the name that has emerged to refer to the open issues between humanities, science, and technology since the mid-nineties. Still, the idea that it is a fundamental feature of the human being to extend the limits

imposed on her by nature or culture and to enhance her traits and capacities is an idea as old as the reflection about human nature itself. Starting from the Italian humanism of the 14th Century, this idea has a strong normative orientation towards education and self-education. This idea spells out that no one ought to remain as nature or the particular social situation has shaped her. Instead, everyone ought to develop beyond their naturally and socially imprinted shape (Pico della Mirandola, 1486/1965).[1]

At the same time, the scope of scientific and technological research has expanded immensely since the early modern period. Since scientific knowledge has been able to translate its laws and principles into devices that "act" on behalf of human beings or enhance them, philosophy has been faced with new issues. During this development, technology and economics have been playing a major role in mediating scientific knowledge (Nida-Rümelin 2005). In the twentieth century, the progress in the field of science and technology has sharpened the debate about human enhancement.

The knowledge that has developed in different fields of science and technology and the linked deployment of potential interventions with possible social relevance has undermined the established order and has given the debate on human improvement an increased magnitude.

The two orders in which theoretically and epistemologically Kant divided the human being are no longer so rigidly separated. According to Kant, man belonged to two orders. The first order is ruled by nature, it is characterized as something "given". On this order human beings are hardly able to make modifications. The second order is the order ruled by human beings. Acting on this order, a human being can change herself and in so doing she is lead both by

---

[1] In this connection, a passage in Pico della Mirandola's *On the Dignity of Man* (1486) comes to mind. He quotes God as saying the following:

> We have given to thee, Adam, no fixed seat, no form of thy very own, no gift peculiarly thine, that thou mayest feel as thine own, have as thine own, possess as thine own the seat, the form, the gifts which thou thyself shalt desire. A limited nature in other creatures is confined within the laws written down by Us. In conformity with thy free judgement, in whose hands I have placed thee, thou art confined by no bounds; and thou wilt fix limits of nature for thyself. I have placed thee at the center of the world, that from there thou mayest more conveniently look around and see whatsoever is in the world. Neither heavenly nor earthly, neither mortal nor immortal have We made thee. Thou, like a judge appointed for being honourable, art the molder and maker of thyself; thou mayest sculpt thyself into whatever shape thou dost prefer. Thou canst grow downward into the lower natures which are brutes. Thou canst again grow upward from thy soul's reason into the higher natures which are divine. (Pico della Mirandola, 1486/1965, pp. 4-5)

her intention to achieve a goal (pragmatic domain) and by moral principles (moral domain), which she recognizes as true[2]. Interestingly, such a moral domain is not doomed to be subjective and arbitrary in scope. Its laws have the same status as the laws of nature already according to the Kantian account.[3]

Since technology and science have allowed us to intervene on the order of nature (that is, both at the level of the human body and at the subpersonal level as in the case of the neural and genetic mechanisms that are modified in order to get cognitive enhancement or life extension) the boundaries between these two orders have become blurred.

The existing framework that produced a divide between "what is given" and "what can be made" was assumed as normatively valid without anything in the way of profound reflection and has been accepted as obvious until recent time. Now it seems insufficient in giving orientation toward the emergence and diffusion of new emerging technologies aiming at human enhancement.

As the demarcation line between the two orders moves, we should rather ask the question whether such interventions raise novel ethical questions that previously did not exist. If these interventions have precedents, which have legitimized their use, then it is clear that there are no new problems.

Indeed, we may say not only that improvement is allowed. It may even be socially accepted, such that all our institutions and social practices presuppose it. The results of these interventions are valued especially when they can be related to the work that is exerted in order to achieve that result.

The question whether new ethical issues are raised becomes controversial if interventions are practiced on the order of nature – in an area, that is, in which humans have had a limited capacity to intervene.

Today it is incumbent upon us to judge quickly and to decide upon new practices that people may gradually want to adopt. We are looking for criteria that help us to decide about issues of enhancement. The first criterion has to do with self-determination. Social sustainability and ethical permissibility, however, are also relevant in this context.

As an example, some conditions that are considered necessary for the attribution of personhood are the same that are invoked as criteria for deciding

---

[2] "Physiological knowledge of the human being, systematically formulated (anthropology), can exist either in a physiological or in a pragmatic point of view. - Physiological knowledge of the human being concerns the investigation of what *nature* makes of the human being; pragmatic, the investigation of what *he* as a free acting being makes of himself, or can and should make of himself" (Kant, 1798, 7, p.119).

[3] "The starry sky above me and the moral law within me" (Kant, 1788, 5, p. 161).

the permissibility of human enhancement interventions. Conditions that need to be met, according to some scholars, in order to ascribe personhood could be considered as the definite ethical boundaries for permissible human enhancement interventions. If self-consciousness, a capacity to weigh reasons for action and a capacity to develop a life plan are preserved, then there are no major concerns that would forbid human enhancement.

The concern that through these new practices human enhancement may furtively promote values other than those we feel bound to (e.g. justice, equity, solidarity, etc.) is at the core of our efforts to identify new criteria for permissible enhancements.

It is not easy to provide ready-made criteria that are appropriate in all cases. The practices of human enhancement call into question existing conceptual frameworks and values, that we have adopted so far, for understanding and regulating the democratic fruition of the results of scientific research. Even if the techniques are applied for therapeutic purposes, they do not just restore a function of the body or the mind, but strengthen and improve human traits and capacities that transcend any disease. Thus a mere bioethics approach turns out to be insufficient. Even a human rights approach does not prove reliable as a criterion because the principles of human rights can hardly be adapted to fit the context of human enhancement. Overall, human enhancement calls for some consideration of our system of knowledge and its articulation in scientific and moral knowledge.

In this scenario it becomes necessary to determine the role of scientific knowledge in the process of understanding and then regulating human enhancement.

Nobody would doubt the fact that we live in an enlightened society in which we rely on scientific knowledge. Rather, the question is the following. Do we believe that scientific knowledge is the sole form of knowledge upon which the work and design of human beings relies? Or do we believe, instead, that there is a right on the part of humans to establish the goals and objectives of their development, which belong to a different domain than logic and the ethos of scientific rationality?

At stake is not only the demand to adopt an interdisciplinary approach - which is also very important to understand and judge with respect to human enhancement. Rather, the settlement of scientific and moral knowledge as well as the demand of an integrative judgment that could harmonize moral and scientific judgment are at stake. It is, rather, an epistemic challenge: to confront the truth of the life-world that can lead individuals to choose a

technology to enhance themselves and the truth of science. This is the reason why the clarification of the relationship between humanities, science and technologies is an integral part of a new analysis of human enhancement that also benefits from the field of metaethics.

3. Desire and Technology

Nowadays, the knowing subject has to deal with distributed knowledge (Verbeek, 2006). The knowledge is distributed in what may be called interacting networks of agents, where the circulation of information brings about the possibility to contextualize the world. In this sense, the human enhancement technologies have expanded systems in which information circulates, becoming more and more complex and ever more autonomous (Floridi, 2013). By virtue of this network structure new technologies are therefore not only forms of instrumental rationality defining the relationship between humans and the world. They also have an impact on how humans feel and think themselves, affecting their emotions and imagination. "Technologies and moralities happen to be indissolubly mingled because, in both cases, the question of the relation of ends and means is profoundly problematized" (Latour, 2002, p. 248). As "desire", "morality", "technology" are not immune to one another, technological tools are related to values not only depending on how they are used, but also on how they are planned and designed. Accordingly, the social and moral significance of technology cannot be reduced to the (right or wrong) use of it. All along technology can lead to an amelioration in human development as well as to bad use or to human hubris (Cerqui, 2002; Coeckelbergh, 2009). Thus, technologies are not merely "neutral" instruments, but expressions of human sensibility and rationality, i.e. of an activity of conscious planning, which is oriented by the goals it pursues.

4. Realism and Normativity (Criticism of Naturalism)

Agency is the centre of the human condition. Our reframing of the debate on human enhancement also focuses on the normativity that is at the core of acts of human enhancement.[4]

---

[4] This has been pointed out by Straub et al. (2012), but also much earlier by Pico della Mirandola (1486/1965).

We may stick to the Kantian distinction mentioned earlier – i.e. the distinction between what nature *does* and what human beings *can* do and *ought* to do. This would suggest that the human domain comprises human capabilities, social mediations and established practices. The human condition reaffirmed the desire to define ends.

In the mainstream discussion on human enhancement, it seems that these elements of the human condition are no longer required. For a variety of historical and conceptual reasons the debate on human enhancement has mainly been developed from a naturalistic stance.

Naturalism is a thesis about the world that may be articulated from an ontological (or metaphysical) or from an epistemological perspective. The former states that all that exists is natural. If, in contrast, the thesis is an epistemological one, then it states something different. It purports that everything we know is known through scientific and technological means. The latter form of naturalism can be called as *scientia mensura*. Sellars provides us with a paradigmatic meaning of this kind of naturalism: "In the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that is not" (Sellars 1956, 173). This means that science and technology are the measure of all possible knowledge. Sympathetic with this position is a kind of moral epiphenomenalism. According to this position, moral knowledge has no truth. This moral epiphenomenalism comes in various forms, e.g. relativism, emotivism, and constructivism, etc. In the case of moral emotivism, moral knowledge is nothing but the expression of moral feelings. This reductive approach and moral emotivism cannot explain our ordinary moral language and reasoning. Neither can it explain the fact that we learn which actions, policies and social settings are ethically good or fair.

If we accept these epistemological conditions, it becomes understandable how it is possible to fall into a kind of naturalism according to which scientific theories are not only true but also the unique source of truth. This position leads to a deflationism in ethics. According to this kind of reductionism, the ethical and social sustainability is downgraded until it coincides with the merely technically feasible.

It is against this naturalistic background that we aim to reframe the debate. In particular, with respect to the quest for normativity, we want to suggest that metaethical arguments that affirm the reality of moral knowledge together with the reality of scientific and technological knowledge can provide a reliable

basis for debate. This would help us to engage in a discussion with the mainstream in the enhancement debate, that is naturalistic and reductionist. In doing so, we hope to reorient the enhancement debate and to put it from its head back on its feet. We hope to redirect some of the attention away from the laboratories and research centres and back to our lifeworld. Furthermore, we hope that the debate will not be given up to purely economic criteria.

Starting from these remarks, it is possible to state that metaethical reflexion may support our aim to make room for a normative-ethical stance in dealing with ethical issues that arise from scientific and technological knowledge which aim at enhance human traits and capacities.

5. Ethics and Law

The debate on human enhancement reveals how important and productive it is to have a broad perspective on different phenomena if the understanding and regulating technology is concerned. Both ethics and law are expressions of normativity that influences the way people interact with each other and engage with their environment. Even if ethics and law are guided by different regulatory ideas, ethics and law are in many cases in line with each other – but not always. Thus, ethical normativity must not be confused with legal normativity. Indeed, in some cases the results of an ethical analysis may be totally out of tune with the legal analysis. For example, in the case of an emergency it may be unlawful to cross the street without regard for approaching traffic. But this does not seem to be unethical. In other cases some result of ethical analysis can be implemented in legal analysis. As a consequence, ethical reflection has a crucial role to play in elaborating legal instruments. In fact, some results of ethical analysis can be very valuable for legal analysis. But some others have just precautionary meaning or are just expressions of moral preferences. These are legally uninteresting, even if they remain ethically relevant.

In any case, the implementation of ethical ideas to regulatory ends raises the question how ethics and law relate to one other (Nida-Rümelin, 2013).

Regarding the relationship between ethics and law some scholars have argued that there is an ongoing process of "ethicalization of law" (Vöneky et al., 2013). They have described this ethicalization as follows:

> One could describe this notion by the phenomenon that legal rules are being supplemented more and more by ethical, non-legal standards / norms. We are

not stating that this is a new phenomenon per se; but we think that we can see it becoming more frequent. There are more and more clauses in legal norms which give ethical (non-legal) norms some validity in a legal order: We can find such "opening clauses" for instance in the Framework Programme of the EU; according to it, "all research shall be carried out in compliance with fundamental ethical principles". Secondly, we can also observe an increase in the establishment of ethics committees on a legal basis. There are more and more areas in public international law, European law or national laws where the decision of an ethics committee is necessary before any action is allowed – for instance in drug trials on human beings. A third area of the ethicalization of law can be found when one looks to non-binding ethical codes of conduct, which are submitted by private organisations (Vöneky *et al.*, 2013, p. xi).

With regard to this complex phenomenon, we believe that this step can be achieved only through interdisciplinary work performed together with legal scholars. Technology becomes one of the forms through which individuals learn to think and desire. Therefore, its regulation has to take two things into account. For one thing, it has to consider the possible retardation of technological developments due to increased research and production costs, which may result from their inherent technological complexity. Furthermore, it has to pay heed to the limited market for technological innovations. The legal system needs to provide solutions for various problems. Regulation may allow the substantial benefits of such kinds of application to reach the market earlier. They may also be available to a larger public. At the same time, it may provide the technical tools and criteria to prevent or, at least, mitigate the (socially) undesired effects of human enhancement.

Ethical normativity depends on people who formulate normative claims or act according to them. At the same time, normativity provides us with true claims about events, objects and relationships that shape the world we live in. In ethics there is no Archimedean point from which we can derive what is right and wrong by following the top-down process of applying the developed principles. According to Julian Nida-Rümelin, both the Kantian categorical imperative and the utilitarian principle should be rejected on that count. Unfortunately, the implementation of these principles cannot give us normative guidance because of the different context-sensitive properties cannot be generalised into a comprehensive ethical theory (see Nida-Rümelin, 2007).

Ethics has different purposes. It is to provide conceptual clarification. (Consider, e.g., the aforementioned issue about the categorization of human

enhancement measures). Another function of ethical analysis is to make our beliefs and values consistent.

6. Politics and Policy

Science and technology is of great economic interest. They have the potential to quickly modify individual and collective lives that no other human artefacts has. Neither law nor religion (which are other human artefacts) have produced so rapid and widespread transformations in human life. These transformations have disseminated into social behaviour without making too much clamour, rather by gripping the desires, the imagination, and the human expectation.

Human enhancement technologies act on emotions as well as on the structure of rationality. In doing so, it is therefore clear that they cannot simply be considered as something to be permitted or forbidden. As human enhancement technologies feed new hopes and create a social demand, make available new tools both for individuals and society, foster threats and concerns, and present risks, they need to be dealt with in a public discussion and not only in academic circles. This can be addressed only in a democratic debate as pointed out by Coenen et al. (2011). Thus philosophical reflection on technology and philosophical reflection on democracy are two joint themes. This is so not just because the technology is an increasingly important tool that allows citizens to access political life, but also for a more significant reason. Only a public, democratic debate can develop policies which allow for a legitimate use of human enhancement technologies that improve the human condition. The debate about human enhancement shows exactly this. That is, it shows how true scientific culture can develop through a public forum that allows for open and free discussion.

7. The Contributions

The volume is opened by Volker Gerhardt's essay. It addresses fundamental philosophical questions and clarifies and deepens the theoretical premises at the core of the philosophy of technology. Gerhardt devotes his attention to the question how philosophy could be disentangled by an allegedly irresolvable contrast between culture and nature. According to him, technology, which plays such an important role for no other species than for humans, paradigmatically shows this evolutionary development of nature, viz. that in order to emerge it has to become something other than itself.

In their contribution, Nikil Mukerji and Julian Nida-Rümelin approach the debate about human enhancement from a general angle in order to make a case for a moderate take on the issue. Based on general conceptual reflections, meta-theoretical considerations and a brief *tour de force* through some of the most important motifs of the enhancement debate, they criticise sweeping conclusions both in favour of and against human enhancement. Instead, they hold that we should consider the ethical issues which arise in the context of enhancement with an open mind and on a case-by-case basis.

The third essay of this issue is by Christopher Coenen, who examines transhumanism from a historical perspective. He investigates the narratives on which it is based and holds that they can explain why our societies are currently so fascinated by the perspective of enhancing human nature. Coenen believes that the "historical interpretative approach" he advocates may, as he puts it, "give rise to a new reflexive stance on current enhancement discourse".

Barbara Henry aims to develop an account of post-human enhancement which makes it possible for us to envision a future society that is made up of human beings, human-machine hybrids and artificial beings which can be viewed as free and equal. In doing that, she distinguishes the idea of the "post-human" from that of the "trans-human". The former is meant to refer to symbols and phenomena different from those that are associated with the "trans-human". She believes that in order to reframe the debate on human enhancement we have to guarantee, as she puts it, the "widest possible conditions of non-hegemonic or expansive conscious contextuality of legislative and decisional systems".

David-Jack Fletcher's essay homes in on transhuman technologies that target and aim to eradicate disabilities. He thinks that the idea of the eradication of disabilities assumes a secular humanist notion of human perfection. And he problematizes the fact that the use of transhuman technologies may lead to "hierarchies of life". Disabled individuals, he fears, may be moved to the bottom of that hierarchy and may even be considered nonhuman. In response, Fletcher offers an alternative view of disability. On that view, "disabled" individuals may not have the same mode of existence as "abled" individuals. But their mode, he holds, may be seen as just as valid as those of the "abled" individuals.

Jan-Christoph Heilinger discusses anthropological arguments about human enhancement. In current debate these arguments are taken to be controversial. Heilinger, however, takes a contrary stance. Based on a

contractualist and pragmatist starting point, he develops an account of anthropological arguments and argues that they can play an important role in determining the ethical status of measures of human enhancement. He emphasises, however, that the content of anthropological arguments is rather minimal and points out that they are ill-suited for supporting a radically sceptical stance on enhancement.

The essay of William Sims Bainbridge addresses the theme of personal identity. In particular, he examines the means of modern communication and computing technologies. He argues that these possess a profound transformative potential and may give us the option to adopt multiple identities, e.g. online avatars, semi-autonomous intelligent agents, etc.

Roberto Mordacci examines three categories of enhancement: cognitive supports and education, neural cognitive enhancers (e.g. drugs) and technological cognitive enhancers (e.g. implants, extended minds). Based on the Parity Principle, he argues that there is no morally relevant difference between the three categories. What we aim at preserving is not the biological status quo of a person's mind. Rather, we aim to preserve personal identities. On that assumption, then, general objections to cognitive enhancement are unsupported as even traditional enhancement measures have their drawbacks and may threaten an agent's autonomy and personal identity.

Anna Gotlib's contribution focuses on virtual realities and the reasons for expanding the idea of human embodiment in order to accommodate them. Gotlib notes that virtual realities play an increasing role for the self-image of human beings. She criticises the prevalent conception of identity-constituting embodiment and argues that virtual environments, such as *Second Life*, can help us to expand our idea of embodiment and to deepen our moral vocabularies of the self.

Filippo Santoni de Sio, Philip Robichaud, Nicole A. Vincent discuss the question when human beings should enhance themselves, focusing particularly on the case of cognitive enhancement. Cognitive enhancement, they argue, is impermissible when it is used in the context of an activity that would lose its entire point due to the enhancement. However, they embrace cognitive enhancement in two sorts of cases: firstly, in what they call "practice-oriented" activities (e.g. of a recreational nature) and, secondly, in certain goal-directed activities (e.g. high-potential jobs), where much is at stake in the way of moral value. Comparatively safe cognitive enhancers, the authors argue, may even be obligatory in the latter case.

Stefan Lorenz Sorgner takes up and criticises Julian Savulescu's suggestion that we may have a moral obligation to increase our children's chance in life by enhancing them. He rejects the "principle of procreative beneficence" which, as Sorgner explains, forms the basis of Savulescu's reasoning and seeks to show that it represents a violent attack on human beings who disagree with it. In conclusion, Sorgner proposes a number of reasons for a principle of procreative autonomy that he takes to be more plausible than Savulescu's principle of procreative beneficence.

Fox Swindells' piece discusses the issue of economic inequality as it relates to human enhancement. As Swindells explains, considerations about economic inequality often lead to a call for the prohibition of enhancement technologies. He believes, however, that a prohibition of that sort would be ineffective in preventing potential harms and may furthermore prevent many positive consequences that enhancements may bring. Swindells also argues against free-market allocations of enhancements as they may lead to unacceptably unequal access. He concludes that governments should step in and provide regulations as well as public funding for enhancement technologies in order to ensure that they are distributed fairly.

After that, we present two case-studies, one by Nathan Van Camp and one of Alberto Pirni. Nathan van Camp surveys the current bioethical and politico-philosophical debate about the so-called 'liberal eugenics'. He argues that the liberal argument for enhancement has internal flaws. However, the liberal anti-enhancement argument, he holds, suffers from the same defects. The latter, he explains, necessarily entails a "preemptive dehumanization" of enhanced life forms, while the liberal argument for enhancement appears to reduce non-enhanced individuals to a "wrongful life". Van Camp concludes, therefore, that the issue of human enhancement cannot be satisfactorily addressed in a liberal conceptual framework. Alberto Pirni offers a survey of the current debate on human enhancement with an eye to its interdisciplinary characteristics. He draws on Aristotle and, in particular, on his theory of justice and fairness that is developed in the *Nicomachean Ethics* to sketch a proposal of his own. It leads up to a synthetic list of possible points of criticism that may apply to the topic.

Last but not least we have Pericle Salvini's piece in which he proposes an alternative framework for human enhancement and illustrates it using the example of robotics technologies. His view is based on the notions of reciprocity and mediation. Salvini argues in favour of the following view. Enhancement, as he sees it, is a result of the way in which technological and

scientific mediation alters the structure of the network of reciprocity that characterises "human presence". Technological mediation, he believes, may turn the "reciprocity of presence" into a unilateral relation that forestalls any form of response.

REFERENCES

Battaglia, F., & Carnevale, A. (2013). *An Ethical and Philosophical Analysis of Human Enhancement.* Internal report of *RoboLaw* project (Available online at: http://www.robolaw.eu/deliverables.htm).

Cerqui, D. (2002). The Future of Humankind in the Era of Human and Computer Hybridisation. An Anthropological Analysis. In *Ethics and Information Technology*, 4 (2), 1–8.

Coeckelbergh, M. (2009). Virtual Moral Agency, Virtual Moral Responsibility: On the Significance of the Appearance, Perception and Performance of Artificial Agents. In: *AI and Society*, 24 (2), 181–189.

Coenen, Ch., Schuijff, M., Smits, M. (2011). The Politics of Human Enhancement and the European Union. In Savulescu, J., ter Meulen, R., Kahane, G. (eds.). *Enhancing Human Capacities.* Blackwell Publishing Ltd, pp. 521-535.

Floridi, L. (2013). *The Ethics of Information.* Oxford: Oxford University Press.

Habermas, J. (2003). *The Future of Human Nature.* Cambridge, MA: Polity Press.

Harris, J. (2007). *Enhancing Evolution: The Ethical Case for Making Better People.* Princeton, N.J: Princeton University Press.

Kant, I. (1788). *Kritik der praktischen Vernunft .* In Kants gesammelte Schriften, Königlich Preußische Akademie der Wissenschaften (ed.), Volume 5, Berlin (1900 -).

Kant, I. (1798). *Anthropologie in pragmatischer Hinsicht.* In Kants gesammelte Schriften, Königlich Preußische Akademie der Wissenschaften (ed.), Volume 7, Berlin (1900 -).

Latour, B. (2002). Morality and Technology. The End of the Means. In: *Theory, Culture & Society* 19(5/6), 247–260.

Nida-Rümelin J. (2005). Wissenschaftsethik. In Nida-Rümelin J. (ed.). *Angewandte Ethik. Die Bereichsethiken und ihre theoretische Fundierung.* Stuttgart: Alfred Kröner Verlag, 834-860.

Nida-Rümelin, J. (2007). Human Biotechnology as an Ethical and Social Challenge. In Knoepffler/ Schipanski/ Sorgner (eds.). *Humanbiotechnology as Social Challenge.* Aldershot, Burlington: Ashgate Publishing, 129–135.

Nida-Rümelin, J. (2013). Recht und Moral. In Vöneky S., Beylage –Haarmann B., Höfelmeier A., Hübler A.-K. (2013). *Ethik und Recht. Die Ethisierung des Rechts/Ethics and Law. The Ethicalization of Law.* Heidelberg: Springer, 3-16.

Pico della Mirandola, G. (1486/1965). *On the Dignity of Man.* Translated by Wallis, C., Miller, P., Carmichael, D., Indianapolis: Bobbs-Merrill.

Savulescu, J. & Kahane, G. (2009). The Moral Obligation to Create Children with the Best Chance of the Best Life. In *Bioethics* 23 (5), 2009, 274-290.

Sellars W. (1956). Empiricism and the philosophy of the mind. In Sellars, W. (ed.), *Science, Perception and Reality*, New York: New York Humanities Press, 1963, 127-96.

Straub, J., Sieben, A., Sabisch-Fechtelpeter, K., (2012) Menschen besser machen In Sieben, A., Sabisch-Fechtelpeter, K., Straub, J. (eds.) *Menschen machen. Die hellen und die dunklen Seiten humanwissenschftlicher Optimierungsprogramme.* Bielefeld: Transcript-Verlag, 27-75.

Verbeek, P. (2006). Acting Artifacts. The Technological Mediation of Action. In: Verbeek, P., & Slob, A. (Eds.). *User Behavior and Technology Development. Shaping Sustainable Relations Between Consumers and Technologies.* Dordrecht: Springer, 53-60.

Vöneky S., Beylage –Haarmann B., Höfelmeier A., Hübler A.-K. (2013). *Ethik und Recht. Die Ethisierung des Rechts/Ethics and Law. The Ethicalization of Law.* Heidelberg: Springer.

Technology as a Medium of Ethics and Culture[*]

Volker Gerhardt [†]
volker.gerhardt@philosophie.hu-berlin.de

ABSTRACT

In the opening of this paper I argue, contrary to a notion commonly found in the humanities and social sciences above all, and ironically yet not any less resolutely for all that, that in fact nature still exists. Subsequently I criticize the arguably more prevalent conviction that culture stands in opposition to nature. In contrast I argue for the thesis that culture is nothing more or less than a specific form of life that can only develop *within* nature. More than this: it can only ever develop as nature; and it can only accomplish this *with* nature. This is shown by examples for the case of technology, which was decisive for the emergence of culture. Every technology can only start by using natural means. Hence we can show that we can only ever understand culture as a part of nature that differentiates itself, like all parts of nature, according to the external and internal conditions of that particular life. Thus culture is one of the countless evolutionary products of nature.

Introduction

The growing self-assertion of the historical and social sciences in the 19th century – which was doubtless also a defensive gesture against the rise of the natural sciences – gave such predominance to the belief in the historicity and sociality of all existence that anyone who appeals to nature in any way seems to be relapsing into mythology. Nature is present everywhere and in everything as the beginning, means and condition; every conceivable end-state of every development, unless it is explicitly supposed to lead to a heavenly realm, is only conceivable as a natural state. And yet argumentative recourse to nature is only

---

[*] This paper is based on a public lecture at Technische Universität München June 17, 2013.
[†] Humboldt Universität zu Berlin, Germany.

accepted in the social-scientific discourses when it promises to produce a social critique.

An ecologist need only condemn the "destruction of nature" and we already forget to ask what "nature" is supposed to mean here. The state of nature in the year 1500? Nature before the first artificial irrigation of Egypt? Or nature before humanity arrived? A sociologist need only deplore the "alienation" of industrial labor and we forget that this critique not only implies an antecedent human nature but calls for its preservation. But how can we wish to preserve something it is forbidden to speak of? And what is the conservation of nature even supposed to mean when all nature is the product of incessant change? When a philosopher such as Jürgen Habermas defends the "life-world" against the instrumental "systems" of modernity, he is making an appeal to a nature that he cannot say anything about, according to the premises of his "critical theory". But identifying this nature is of primary importance if we want to make any meaningful use of the concept of evolution. Moreover, the evolutionary process itself, which no-one can seriously doubt anymore, shows clearly that nature has a history too.

It is typical in the justifications of the cultural and social sciences to only recognize those factors that are themselves socially accessible. In political contexts this sort of approach is called "imperial". "Imperial" means that a power wishes to control the conditions of its recognition by other powers; and this is precisely what we achieve by suppressing the natural conditions of social, cultural or political phenomena. We think that "nature" is no longer pertinent as soon as we start debating social, political or cultural phenomena. Hence, to take just one example I can think of, nature practically never receives any mention in political theories – unless, that is, we think ecologically and call for its "conservation".

What should we make of this self-restriction to allegedly self-made causes? In light of the existence of the natural sciences, this hardly requires comment. What should we say, after all, when it should be obvious, if we pay even just a little attention to our concepts, that every meaningful reference to social phenomena presupposes at least their delimitation from nature? This delimitation only serves to secure the discipline methodologically.

In point of fact, nature is present in every single process and in every thing, since without nature there would be neither facts nor data. And when we start looking at the composition of individual occurrences, it is hard to find anything

that lacks some proportion of nature. Hence nature is at work when it comes to action, speech and thought as well.

1. Culture as the Differentiation of Nature

A closer look shows us that every possible description of social circumstances must involve an awareness of countless natural aspects as well if we are to maintain even a minimum of faithfulness to reality. Try to imagine, for example, a sociological analysis of actions that did not presuppose corporeally present individuals and did not presume the existence of either objects or motives. Try to understand what the meaning of a statement could be that was independent of all needs and expectations as well as all temporal conditions. What could we think of conventions, institutions or religions that did not need any date in history, since they did not bear any reference to a geographical location or a set of people or an intention or desire that anyone might have had at any time?

Hence it is not surprising that nature has not truly disappeared even from modern political theories. It just does not get raised as a topic. The concept of a "state of nature" is an exception, since it is meant to indicate a limiting condition of the state. The "state of nature" after all is something to be *left behind*. In other cases, such as when we talk about birth, illness or death, or of the individual, the few and the many, of lack and surplus, threat and punishment, or life and corporeality – or the conditions of a crisis – we need only recall what we are talking about and the presence of nature underlying everything becomes palpable.

Moreover it does not rob the spheres of history and society of any of their dignity to admit that they are bound up in natural circumstances. We need only trace the allegedly purely social-scientific phenomena back a few historical lengths -- to inquire into the medieval family, the city in the Homeric age, trade before the invention of writing, or stone-age technology – and it is clear how much they are embedded in nature. We see that nature is not just externally bound to culture, that it shapes even the very internal dynamic of culture. Culture gains its particular character not by removing itself from nature, but by its *specific differentiation* of itself from nature, *with* nature and *within* nature. It is a *self-differentiation* that allows nature to enter into an opposition to itself and incessantly make more space for its own potency.

If we can see nature itself as a kind of living being, we can also say that with society, culture and politics it has formed organs that are especially useful to nature in that they can always oppose nature. Yet even this fundamental opposition is still always part of nature, which maintains and develops itself in oppositions.

2. Nature as Outer and Inner Constitution

An example might help to clarify how nature determines more than just our mere existence: we can only sculpt stone by perceiving and recognizing the material qualities it has "by nature". This knowledge of the nature of the material must also be reflected in the tool, which must be suited to the hardness and brittleness of the stone. If the tool is made by working some other material, then the procedure to make the artificial material, such as iron, must be suited to the raw material, such as ore. If the stone to be sculpted is soft, a piece of granite might suffice as a tool. All work on any material makes use of materials that need to recognized and understood in turn, even if we change them.

In making tools we use parts of nature against other parts of nature, and do so by means of the natural constitution of our bodies and the abilities we master through our bodies. Since the specific capacities of the human body only emerged in the course of evolution, there is no reason not to call the products of an additional cultural evolution "nature" as well. Otherwise we would have to dispense with every use of the concept of nature and could at best speak of an infinite series of cultural stages. For in nature everything is already "made" and thus is always already the consequence of a mechanical and organic self-cultivation of nature.

It is also clear that this process of humans working the stone takes place within nature. The whole affair might occur inside of a cave or outside, in the rain or the sun, leisurely or under great time pressure: humans always need air to breathe, enough space for their movements and a fixed stance that allows them to exert force. Humans need powers that, as an expression of their nature, they must develop, practice and renew by nature, in order to even be conscious of them as *their* powers.

The human worker needs to understand not just the available material but also the conditions of their own physical constitution – if they are to be successful in what they do. We can feel, grasp, press and hit with our hands but

we cannot see, hear or smell with them. For some activities one finger is enough, others require the swing of an arm or the exertion of the whole body.

In all we hope to achieve with our hands we count on the success of our bodily movement, which we (as humans) coordinate as a whole in our succession of movements. But then no matter how little some of us might know ourselves we must always understand our own nature. Only by this confidence in our own bodies can we achieve an effect *as* nature, *in* nature, and *with* means won *from* nature, an effect which can be assessed as the evidence of cultural achievement.

It is of course not enough to see humans as composed of nature just by virtue of our bodies. Why do we work the stone at all? Because, we suppose, the sculpted stone is known or expected to bring us some advantages that are not hard for the user to recognize. The hand-axe can help to work fibers, fur or wood, to dismember fruit or a killed animal or to defend against enemies; and intentions that stand in relation to the nature of humanity are operative in all these contexts: because we do not grow our own fur, we need clothes and shelter; since we lack fangs and claws, we need tools to help break down our food; nor are we exempted from the competition between species and individuals at work in nature as a whole.

Even the very logic of need and necessity rests on nature. We have and keep "natural" enemies that we need to resist. For this reason we take up arms, and it is only because we wish to hold our ground *in* nature, *with* nature, *as* nature, that at the end of a long cultural development we practice politics as well.

3. The Dense Context of Nature

In our own self-facilitated process of development we do not just adapt to natural circumstances. We identify with them as far as we can; other circumstances we fend off as alien, adverse or hostile. *As* nature, *in* nature and *with* nature we have always had to assert ourselves against certain parts of nature. Even in shooing a fly away from our food we are using *ourselves* as nature (as a needy and irritable creature) *in* nature (in an environment we share with flies) *with* nature (with our hand) *against* nature (the insect) to protect one of our natural needs. No matter how much effort and skill went into preparing the meal: as food it is part of nature, which must become us as parts of nature and be protected from that other part of nature we recognize as the fly. There

could hardly be any denser locus of effects than that in which humans stand with their fellow humans, animals, plants and other things.

In this context humans must differentiate nature, *for*, *with* and *within* themselves, into that which is productive, counterproductive, or indifferent for their goals. This also occurs naturally: we seek to assert ourselves *as* nature *in* nature and *with* nature but also always *against* nature. We are natural creatures in our perceptions, feelings and drives all the way to the most sublime motives of our understanding of self and world, and remain such, even if in the course of our natural self-assertion we alter our abilities and capacities so much that *culture* arises.

The more specialized our abilities and capacities become and the more they depend on externally perceivable techniques, the more sharply culture stands out. Yet even with this culture we do not suddenly or gradually jump outside of nature! Rather in this elaborated culture we are still necessarily bound to nature. Hence there is no breaking or withdrawing from nature. To the contrary, culture is the expression of the productive participation of humans in nature and its history.

This participation also includes the mechanical utilization of nature, since this stems from our needs and satisfies them. Even if in an individual case it might seem strange and even detrimental in some respects, it is only possible as the result of a cultural development, hence: it is only possible organically. What else could culture produce except figurations of life for the purpose of life? The use of fire, for example, began the course of human culture around 2.5 to 3 million years ago (Schrenk, 2003).

Hence technology, which has probably played the decisive role in the emergence of culture, should be seen as the essential moment of nature's continuing self-organization in culture. Successful technology is the adaptive intervention in nature as a result of which culture emerges. We can see technology as the realization of an active participation in nature. At the same time there are good reasons to reserve the term "participation" for the self-conscious and free involvement in a process based on individuals' self-determination. And yet we can see everything we are familiar with as technological players in nature as paradigmatic of the demands that morality and politics make on every independent person. Hence we are not wrong to speak of our use of technology in terms of "participation", which prevents us from inappropriately rebuffing technology as mere mastery over nature.

Humanity as the purported *maître et possesseur de la nature* is a pious hope, or an impious hope – depending on how you see it. But either way it is a *hope*. For however much he commands parts of nature, the alleged commander remains bound to nature. In our claim to mastery we are following a natural impulse. This holds of all cultural activities, particularly politics, where we find it particularly difficult to discipline the forces of our nature that politics has unleashed.

Hence despite their air of autonomy, morality and politics cannot go beyond *participation in nature*, neither by efficient anthropotechnological self-discipline nor by means of the institutional technologies of legal systems, bureaucracies or the military. Nonetheless, analogously to individual self-consciousness, they can take on an "autonomous" position towards other societal processes.

4. The Problem of Conscious Internal Space

Consciousness is also a real natural quantity. *Idealism, solipsism* and *epistemic skepticism* are the consequences of starting with just individual consciousness and its current contents. If we take consciousness to be just what one has *oneself,* and can only attribute to others hypothetically, then the subjectivity of consciousness is a dead-end that cannot take us to the world or other minds. When in reflexive self-justification we look for a "ground in consciousness" that can bear up everything else, we only sink deeper and deeper into an inescapable rut – as the work of Dieter Henrich made clear (see Henrich, 1991). There is no route from the subjectivity of pure self-referentiality to the categorical space of consciousness that includes us *and* the world.

It seems the only possible solution would involve divine assistance, a divinity that could bring together the world and individual consciousness in one sweeping mind. This is how Plato claimed to give the ideas efficacious reality; this is how Descartes bridged the gap between *cogito* and *res extensa*; and ultimately Hegel found himself mediating concept and reality in the apotheosis of historical becoming. Hegel was skeptical of the solution Kant had introduced, that of overcoming the self-isolation of consciousness with a logical *inference* from the inner space of consciousness to a necessarily correlative external world; thus with an ingenious transfer of the intellectual process of dialectic into the real historical process, he had recourse to a concept of mind that would ultimately prove to be divine.

Hegel was not mistaken in his critique of Kant. For Kant's "refutation of idealism", which seems irrefutable at first glance, relies on the analytical correlation between "inner" and "outer" that rests on the achievement of consciousness proving something to *itself*, without providing any factual evidence that this outer world *really* exists. The outer world must be *thought*, but no argument, no matter how solid, can guarantee that it actually exists. Thus, Hegel concluded, Kant did not succeed in really connecting the space of consciousness with the actual world.

The solution proposed in the 20th century, namely that of locating a bridge to the outer world in the workings of consciousness by means of the intentions inscribed within consciousness, likewise falls prey to the same suspicion that the argument is meant to *divulge* something. We can see this empirically in the unabated currency of epistemic skepticism. But we can also see it quite clearly in the insufficiency of both the evidence and the arguments for the objective achievements of individual consciousness. For in every case, regardless of whether we rely on Husserl's notion of intention or the sense of intention much discussed in the wake of Anscombe, we find consciousness itself drawing an inference to an achievement of consciousness that transcends the consciousness itself. Proponents of these arguments point to structures of understanding or speech in order to infer from these to the reality of that which is meant to be understood or spoken.

5. Consciousness as a Social Organ

We do not need to take the problematic route of inference if we can show that consciousness *is originally directed at others for others*. And this can be seen, as I recently tried to show (see Gerhardt, 2007 and 2012) in the role and function of self-consciousness: in everything it achieves it is founded on *the existence of the world* and in everything it effects it performs *functions of mediation in this world*.

Whether we assume the possibility of pure description or knowledge, whether we see it as attentiveness or a self-satisfied hermeticism, whether it concerns intentions to communicate or act or to retain something in memory – in all these cases the achievements of consciousness are not first and foremost in a *conceptual* relation to the world. Consciousness finds itself in a *real causal context* of physical relations, of which the body is a part and which expands

under the complex conditions of life to include social and cultural unities – without ceasing to be nature.

Consciousness takes part in this organization of materials and forces. As consciousness it is an *organ of the organism*, and in its achievements it is originally and irrevocably *individual*, like the organism itself. But like the organism, it does not just take part in the organization of life, it also *represents* it – even if in a new way, namely *technologically* and at the same time *culturally*. For consciousness is able in a quite unique way to make itself an organ of a community of living creatures and understand itself as a part of this.

In all the achievements of its conceptual organization the individual consciousness is necessarily universal, since it is capable of grasping and sorting a *particular* thing (such as a hand-axe) or an *individual* event (such as the emergence of a danger) such that it can be not just *conceived* but also *communicated* in its general characteristic (as "useful", "dangerous", "useless" or "indifferent"). Thus consciousness becomes the *representative* not just of *what is known* but of *what knows* and through *communication* proves to be the registering and coordinating *sensorium of the group* to which the individual belongs. And wherever this happens the individual acting with consciousness becomes an *instance* and a *representative* of the *social bond* in which it always understands *itself* as well.

Thus in its achievements consciousness is never merely the organ of the individual that understands itself (and its world) in consciousness; in its function it always becomes the *organ of the group* in which it was able to arise, in which it articulates itself, in which it can be understood and effective. This can go so far that in this social role it can even take sides against the organism that bears it. Everyone who conforms to the habits of his group, who seeks to satisfy the claims of his family or the duties of his office, or the demands of a certain task; everyone who sacrifices something for a sport or for art or science, who endorses restrictions for the sake of the environment or stands up for basic rights in the name of humanity – and who really means what he says: everyone who commits himself in this way to any issue or task never uses his consciousness merely as the organ of his organism, never merely as the advocate of his own person, but always also as a *representational instance* of the natural and social unities on behalf of which he acts.

6. Consciousness – Independence in Context

The question is, if we view the organizational work of our consciousness as the element and instrument of organization, what sort of organization must this be? At first we might assume it must be the organization of one's own body, which maintains its own life with the help of consciousness. The body relies on an exchange of matter with the surrounding nature. To come into being, develop and reproduce it requires cooperation with others like itself. It moves around in its environment by using it for its purposes: by means of its senses it learns how to direct itself in a larger radius, with a greater use of means and in coordination with others.

Consciousness is useful to us in all this – no matter how many examples we find where it gets in the way of functional expectations: it helps increase our mobility and flexibility in various environments, adapt to unfamiliar and even hostile living conditions, and boosts the utility of natural resources. It is bound to (physical) conditions and (psychic) affects (how could it be otherwise?), yet it is precisely in these bonds that it serves its independent function as the instantiation of intelligence from which we can expect clarifying insights and knowledge.

Thus we are the only creatures who cook, fry or smoke our food by means of our consciousness. This considerably expands the range of foods we can enjoy; many things that were indigestible can now be eaten without any negative consequences; foods that spoil quickly can be kept longer. The pallet of foodstuffs processed by the human organism becomes richer and more constant. Self-cultivating activity expands the natural basis of human life.

And all this can be made *instantly communicable* if need be, so that it is not just effective in that moment, but can continue to have an effect! Consciousness is a social organ, close to a *cultural institution*, that enables *rapid learning* and allows us to learn generationally from *conceived experiences* and from *dangers and solutions we only imagine* (see Gerhardt, 1999, 2000). Thus the current moment becomes the *present*, which can be made the point of departure for *future* and *past*. In this way time itself becomes the imagination of an extension from which various spaces of action open up.

This has all been made culturally productive in a real and astonishingly short span of time that has no parallel in the history of evolution. What other form of life has managed to produce so many changes in its environment at such an accelerated pace? Which other animal has taken up so many other

earthly procedures and turned them to use for its own purposes? I will mention just fire here.

And just as we trust in the effects of fire, we later begin to trust in our tools and get used to the animals we tame long before we become sedentary. Animal husbandry and agriculture reveal more clearly how much observation of nature, conscious work and self-discipline are necessary in order to make possible this cultivated life-world (which was highly cultivated, though it seems relatively natural to us today). And when we are successful, the consciousness we have invested in our own culture meets us halfway and we learn to orient ourselves towards the changes we have made in our living environment.

Hence we can say: in the circumstances of human culture consciousness stands out as a *real causal factor*. Technology and the actions it makes possible create new living conditions that are individually and collectively taken up in the forms of human culture and thus made in turn into conditions of the developing human life. They thus become part of nature, and there is no need for any inference on the part of consciousness from the inner to the outer, a "real" external world. Strictly speaking the concept of the external world becomes obsolete, since there is only the one thoroughly efficacious world to which humanity acquiesces without remainder; even if consciousness gives us the possibility of distancing ourselves from nature procedurally. This is the distance that every living creature creates for itself with its organic self-differentiation from its environment, and that can be expanded by the social experience of our difference from other individuals and from groups we experience as "foreign".

Yet independently of the functional distinction between inner and outer, it is necessary to see *consciousness* as an effective element in the process of the self-developing nature of humanity. Due to time limitations I will restrict myself in the following to two brief pointers in this direction.[1]

7. The Public Interaction between Mind and Reason

Perhaps the most important step in the development towards humanity, even if it could not be seen from outside, was the *institutionalization* and *instrumentalization* of consciousness: with the use of tools and the division of labor, the need to communicate about matters of fact we can reliably and predictably

---

[1] For more on these points, see Gerhardt 2007, Gerhardt 2011 and Gerhardt 2012b.

count on becomes increasingly urgent; this makes communication itself and its means into an *object*, something we can negotiate over like any other objective matter. Whatever we communicate such that it can be understood independently of the given situation and the particular agents involved takes on an existence of its own relative to the speaker and a meaning corresponding to the spoken word. The communication gains a *sense* attaching solely to its expression that is no longer dependent on the presence of a certain speaker.

In this way the original achievement of individuals, manifested in a perception, an insight or a memory (which rests on countless social presuppositions by way of language acquisition, grammar and logic) becomes an institution that individuals must conform to if they wish to use it for their interests. It is not just the medium that stands out; rather we perceive *meaning as such*. It appears to us as a something that we can speak about, as if it were an object like any other.

What is new here is that consciousness and its contents make themselves *independent*. In consequence it becomes possible to speak of "consciousness", of "mind", or of "reason" like particular organs or substances or instruments. And when the contents of consciousness become "meaning" and "sense", "representations" or "thoughts" or "ideas", this is due to the *autonomization of that which is meant* by them.

When that which enables substantial communication attains to self-standing independence, it meets all the conditions of *institutionalization*. In this way consciousness becomes an organization comprising all people who communicate under its conditions. We can thus see it as "mind", as a collective consciousness that comprises everything that can be conceived or coherently thought, without being bound to an individual consciousness.

And in the same way we can describe the achievements of consciousness as "reason", which in its functions of understanding, justifying, judging and deciding formulates means-ends relations and thus aims at the justification and critique of intentions, goals, and suitable forms of action.

8. The Meaning of Signs

The material consolidation of the semantic sphere, which I have only outlined functionally here, occurs, I assume, in the explicit *use of signs*. After the *use of tools* based on the actual operations of the hand, the equally instrumental *use of signs* (in a broad sense that includes all intentional marks as well as pictures

and symbols) is the most important developmental step in the history of human culture: I take it to be the *technical innovation of culture* par excellence that we no longer rely on the *presence of our bodily movements* to communicate with others like ourselves, and instead can make use of *objects produced with tools* that now not only produce physically efficacious *consequences* independently of the presence of the agent (such as traps) but that also produce *signs* evoking *mere meaning*. And just as we use the body, particularly arms, legs and hands, in the manner of tools in order to create and use further tools, we now all of a sudden also use our mouth, tongue and lips in artificial ways to generate self-standing phonetic signs in the social space.

It should be noted that the voicing of articulated speech follows the model of our instrumental dealings with devices: "Language is an instrument", as Wittgenstein says.[2] We learned to use language after having already understood how to deal with tools. And with language we began to shape our lives from the perspective of an open world and with the support of society and of knowledge. The worldwide migrations of early humanity give us a sense of the real openness of the world in our human development. I suspect I do not need to emphasize that point here.

Yet technology might have played the largest role close to home: clay tablets with images, slabs adorned with signs, painted temple walls, writing on animal skins or tortoise shell prevailed as the physical bearers of meaning. They could not produce sense without knowledgeable readers, but they retained their message independently of speaker and hearer. Since their sense renewed its presence with every transcription or telling – and in certain circumstances could even be reproduced with unequivocal conceptual precision – they offered a message that could even exist independently of the particular medium of their bearer. Thus media are necessary to fix a sense conveyed by signs; yet the *replaceability* of the media is one of the decisive reasons why we can see meaning as something that can dispense with any material substance.

Our success in gaining mastery over nature, producing our own tools, practicing techniques and establishing our own customs most likely paved the way for the use of our own bodies to produce meaningful signs. This also includes the millennia of practice producing material for objects that, once

---

[2] Wittgenstein (1953), § 569.

their utility was recognized, never had to be mere tools but could also serve as valuables, cult objects and jewelry.

Moreover we only had to make the meaning invested in them explicit to give them binding force. In this case it could have been the technological devices themselves, particularly in their cultic use, that acted as signatures, paving the way for the invention of writing. We might see such early beginnings of writing in cave painting.

9. The Role of Writing

In writing we find a bearer of meaning that is detached from the speaking individual and that gives expression to the *mind* already operative in all technological and social achievements of humanity. It can work independently of the human. *In this way the self-discipline of the natural creatures that we are, made possible by technology, becomes the starting condition of the genesis of the human mind.*

Whatever the historical process might be: after the invention of *written signs*, after we established public *places of assembly* and set up zones of *mastery and control*, and, as it seems, art helped to so completely fill our living spaces with our self-created meaning it seemed as if they must be comprehensible to higher beings, language must have gained in precision. With language all those who wished to make their mark through the spoken word, the prophets and priests, singers and wise men, actors and raconteurs, were able to expand their medial autonomy.

Yet even beforehand writing had given individuals and whole societies the capacity of *precise* memory. It was possible to write down who owned exactly what piece of land, how much tribute someone had to pay, and under what conditions pasturage rights or a wife's property could be transferred.

Writing allowed the meaning of a process conjointly understood in that moment to be fixed for the future as well. Thus the space of meaning seemed frozen in time, whereas it was in fact unceasingly traversed by new events and conventions that created new states of affairs.

It is almost as if we could see the material consolidation of the semantic sphere occurring in front of us if we imagine how writing made it possible to exactly fix an event over great distances and several generations. The original copy of a purchase contract, the laws of a ruler carved into stone or often into wood, a calendar of kings and their deeds or the records of the movements of

the stars seemed to set down in perpetuity what had happened (for all future readers and hearers) through a one-time act.

All of this generally had to be carried out by individual writers. But it acted on the future in which it was *in principle accessible* to everyone who happened upon it and could read it. This didn't necessarily have to include only those who could read. For the creation of a seemingly self-standing space of meaning it was enough for the laws, contracts and observed data to be accessible *per se*. They could then be conveyed to others orally. And since it can always be deciphered again and start new lines of transmission through reports and stories, writing can be seen as the bridge to what we have to see as the space of *remembering*, of *understanding*, of *thought* and *mind*. We can see that it is more than just "writing" or "language" in the fact that we can understand it as something "public".

Conclusion

Thus I can conclude by speaking *with* Wittgenstein *against* Wittgenstein: "Where our language suggests a body and there is none: there, we should like to say, is a mind."[3] This is exactly right! But since "suspecting" may not get *spoken* but must at least be *thought*, we can say, correcting a misunderstanding that continues to afflict us even today, that in this thought we have at our disposal not just *language* but also *mind*.

REFERENCES

Gerhardt, V. (1999). *Selbstbestimmung – Das Prinzip der Individualität.* Stuttgart: Reclam.

Gerhardt, V. (2000). *Individualität – Das Element der Welt.* Munich: C.H.Beck.

Gerhardt, V. (2007). *Partizipation – Das Prinzip der Politik.* Munich: C.H.Beck.

Gerhardt, V. (2011). Monadologie des Leibes. Leib, Selbst und Ich in Nietzsches Zarathustra. In Gerhardt, V. (ed.), *Die Funken des freien Geistes – Neuere Aufsätze zu Nietzsches Philosophie der Zukunft,* Berlin/New York: de Gruyter, pp. 161– 232.

[3] Ibid. § 36.

Gerhardt, V. (2012a). *Öffentlichkeit – Die politische Form des Bewusstseins.* Munich: C.H.Beck.

Gerhardt, V. (2012b). Humanismus als Naturalismus – Zur Kritik an Julian Nida-Rümelins Entgegensetzung von Natur und Freiheit. In Sturma, D. (ed.), *Vernunft und Freiheit – Zur praktischen Philosophie von Julian Nida-Rümelin.* Berlin/New York: de Gruyter, pp. 201-226.

Heinrich, D. (1992). *Der Grund im Bewußtsein. Untersuchungen zu Hölderlins Denken (1794 – 1795).* Stuttgart: Klett-Cotta.

Schrenk, F. (2003). *Die Frühzeit des Menschen – Der Weg zum Homo Sapiens.* Munich: C.H.Beck.

Wittenstein, L. (1953). *Philosophical Investigations.* Oxford: Basil Blackwell.

Towards a Moderate Stance on Human Enhancement[*]

Nikil Mukerji[†]
nikil.mukerji@lrz.uni-muenchen.de

Julian Nida-Rümelin[‡]
julian.nida-ruemelin@lrz.uni-muenchen.de

ABSTRACT

In this essay, we argue against radical ethical views about human enhancement that either dismiss or endorse it *tout court*. Instead, we advocate the moderate stance that issues of enhancement should be examined with an open mind and on a case-by-case basis. To make this view plausible, we offer three reasons. The first lies in the fact that it is difficult to delineate enhancement conceptually, which makes it hard to argue for general ethical conclusions about it. The second is that an appropriate view of the edifice of moral theory suggests that tenable moral judgements about human enhancement are the result of a careful consideration of the pros and cons that attach to the use of a specific enhancement technology. Lastly, we show that important normative factors in the enhancement debate can be used both in arguments for and in arguments against enhancement. The bottom line of our discussion is that we should treat issues of human enhancement like we do any other ethical issue, *viz.* by weighing up the reasons pro and con.

Introduction

In connection with human enhancement, many important ethical issues arise. Granted, certain technologies that are currently being discussed still belong in the realm of science fiction. Some of them may never be technologically feasible. Others, however, will sooner or later become very real and available. And once they are, we as a society need to be prepared. We need to have a clear comprehension of the ethical dimensions of these technologies before we can enshrine into law the rules that shall constrain their use. The current ethical debate about human enhancement is marked by a sharp dispute between enthusiasts and sceptics. In this essay, we try to steer a course between these two extremes. We want to make a case for a moderate stance, which avoids both overexcitement and undue scepticism. In our view, the ethical challenges of human enhancement should be approached like any ethical issue, *viz.* with an open mind and on a case-by-case basis.

We proceed in three steps. In the first section, we start with a few well-known remarks on the notion of human enhancement. In doing that, we seek to cast doubt on the idea that the enhancing use of biomedical technologies can be clearly delineated, e.g., from their therapeutic use. This lack of conceptual clarity suggests, in turn, that general claims about enhancement are hard to support. In the second section, we clarify, from a meta-theoretical stance, how one should go about discussing the ethical issues that enhancement technologies give rise to. We oppose rationalism and favour a broadly pragmatist conception, which takes our lifeworld seriously. The latter view suggests that the right way to approach the issue of human enhancement is to consider all normative considerations that may reasonably be seen as relevant to the issue at hand and to weigh them up in order to arrive at an all-things-considered moral verdict. This idea, too, casts doubt on any radical position. In the third section, we briefly consider a number of normative considerations that have been employed in the enhancement debate. We explain that all of them can be used in pro-enhancement and anti-enhancement arguments. This suggests that we had better examine each issue separately, carefully and with an eye to all relevant facts.

1. Human Enhancement – The Very Idea

The idea of human enhancement is not new. It has been a constant concern of scientists, philosophers, public intellectuals and politicians throughout the

course of history. In this connection, of course, Nietzsche's famous idea of the "Übermensch" comes to mind.[1] But there are various literature sources which attest that the notion of human enhancement dates back much farther than the 19th century. It can even be found in one of the earliest pieces of literature, the Sumerian *Epic of Gilgamesh*, in which the king Gilgamesh sets out to find a substance that can turn him into an immortal. Ever since, a fair share of writers has spun tales about the fountain of youth. And a number of explorers – most famously, perhaps, the Spanish Marquess Juan Ponce de León – have embarked on a quest to actually find it.[2] Their attempts have, of course, been to no avail. Science, however, has progressed to a point where it has become possible to enhance human beings in ways similar to the old speculations. And a number of incredible innovations may be just around the corner. Though significant ethical issues attach to the possible deployment of these technological advances, advocates of human enhancement paint an auspicious vision of a transhumanist utopia (e.g., Bostrom 2003).

But what is "human enhancement" anyway? As it turns out, it is quite hard to pin down the precise meaning of that term. There is a weak notion, according to which humans have been engaging in enhancement for a rather long time (see, e.g., Allhoff et al. 2009). Our ancestors developed all sorts of gadgets, e.g. bow and arrow, which helped them to become more effective hunters. The use of binoculars has helped them to improve their vision manifold. And through the use of pen and paper humans have vastly extended their recollection and reasoning capacities. Though these examples are undoubtedly forms of human enhancement in a weak sense, they are not enhancements of the problematic type that is at issue in the current debate. What, then, do ethicists talk about when they discuss human enhancement?

It seems that when ethicists use the term "human enhancement" they predominantly intend by it the more drastic types of intervention into the human body that use the means of modern biotechnology and biomedicine. But they do not mean all of them. Self-declared enhancement sceptics often insist that they are welcoming of the means of biotechnology when they are used purely for the sake of medical treatment (e.g. Kass, 2001 and 2003; Sandel, 2007). They oppose their use only in cases where they are deployed in order to give a person "supernormal" abilities.

---

[1] It is, however, contested whether Nietzsche should be seen as a precursor of today's transhumanists who advocates. See the exchange between Bostrom (2005a), Sorgner (2009) and Hauskeller (2010).
[2] Bostrom (2005a) offers an instructive overview over the history of transhumanism.

The enhancement/treatment distinction is often illustrated using the example of Retalin. Retalin is a psychoactive drug that was first developed in the 1960s in order to treat patients with attention deficit hyperactivity disorder (or ADHD, for short). Nowadays it is widely used by "normal" college students (particularly in the USA) to improve their ability to focus and to work longer hours.[3] Sceptics about enhancement may, then, condone (or, perhaps, even welcome) the initial therapeutic use of Retalin. But they would be critical of its second, enhancing use.

Prima facie the enhancement/treatment distinction seems to be a good way to draw the line between enhancing uses and non-enhancing uses of biotechnology. But well-known problems are associated with this distinction. It obviously presupposes the specification of a point of reference that defines "normality". In principle, there are two options. One way to specify the reference point is with regard to a given individual. Call this the *intrapersonal rendering* of the treatment/enhancement distinction. Suppose that a given human individual naturally (i.e. without intervention) possesses a characteristic C to the extent x. x is, then, her "normal" level of C. Any intervention which would increase her level of C beyond x and to the point of, say, $x+e$ would then be viewed as an enhancement. In contrast, if the individual had lost C to the extent i (perhaps due to injury) and an intervention would restore C to its natural level x ($=x-i+i$), then that invention would be considered a medical treatment. Now the problem with the intrapersonal rendering of the normal reference point is that every treatment of an inborn condition (e.g. ADHD) is considered an enhancement although, intuitively, it should count as the medical treatment of an inborn condition.[4] The *interpersonal rendering* of the treatment/enhancement distinction gives rise to a different problem. If we specify the normal expression of C as lying within a certain range of deviation from some average level x in a given population, then we have to choose an arbitrary cut-off point. E.g., mental retardation is commonly defined as having an IQ of 70 or below (i.e. two standard deviations below the statistical mean). On the interpersonal rendering of the treatment/enhancement distinction, we would have to say, therefore, that the medication of a person with an IQ of 69 is

---

[3] In connection with ADHD, there is, of course, the worry that "the borderline is shifting to include more people in disease and disorder categories." (Schermer, 2007, p.34)

[4] As Bostrom and Roache (2008) point out it also leads to the counter-intuitive conclusion that an unenhanced individual, who has a high natural level of C, can possess C to a greater extent than an enhanced individual, who happens to have a low natural expression of C.

a treatment, while the medication of a person with an IQ of 71 is an enhancement. This, of course, is unsatisfactory. A sharp treatment/enhancement distinction seems, thus, unwarranted.

There are, to be sure, various other attempts to draw a sharp line between enhancing and non-enhancing uses of biotechnology. Given the scope of this essay, we cannot address them here. Let us briefly remark, however, that all of them seem to us rather unsatisfactory.[5] If we are right about this, then this is, of course, bad news for anyone who seeks to argue for a sweeping conclusion in favour of or against human enhancement. Does this mean, then, that issues of human enhancement are intractable? Of course, it does not. It is still possible to proceed on a case-by-case basis and to consider the reasons that speak for and against a specific enhancement measure. This procedure does not require a definition of enhancement. For if we only look at the issue at hand and the reasons pro and con, the question whether we are talking about an instance of human enhancement becomes irrelevant.

2. Human Enhancement and Moral Theory

In this section, we would like to strengthen our case further by making a few general remarks about the edifice of moral theory. It is clear that when we debate an ethical issue, such as human enhancement, we first need to agree – albeit implicitly – which view of moral justification we want to follow. For if we fail to do that, we fail to agree on the criteria that a tenable answer to our moral problem has to fulfil. This, in turn, will make it hard to reach agreement on the issue at hand. There are, broadly speaking, two views of moral justification. The first is based on the idea that all answers to moral problems follow – directly or indirectly – from a first principle that is absolutely certain and provides an unshakable foundation for knowledge (lat.: *fundamentum inconcussum*). This idea is often associated with the work of René Descartes. In his *First Meditation*, Descartes writes that it had occurred to him

> that at some stage in my life the whole structure [of opinions, NM & JNR] would have to be utterly demolished, and that I should have to begin again from the bottom up (...). (Descartes, 1641/2008, p. 17)

---

[5] Of course, it is possible to define enhancement in a strategic way, such that it becomes possible to support general conclusions. But this seems to require a gimmicky definition that would presumably not capture what we intuitive understand by the term "enhancement".

Shortly after that, Descartes remarks that "up to now whatever I have accepted as fully true I have learned either from or by means of the senses" (Ibid., 18). His senses, however, have, on occasion, deceived him. And thus, he reasons, sensory data cannot provide an indubitable foundation for all of knowledge. In his view, only a rationalist position is tenable because only *a priori* truths, i.e. truths of *reason*, cannot be doubted. The one on which he chooses to rebuilt his system is, of course, the truth that he exists, which can be known *a priori* from the indubitable fact that he is thinking, as Descartes' famous *cogito* argument has it. It provides the unshakable foundation for all knowledge. At any rate, so argues Descartes.

In ethics, the prime example of the Cartesian rationalist position is that of contemporary utilitarianism, as it is held, e.g., by Hare (1981), Singer (1979/1993) and Smart (1973). The utilitarian system contains one fundamental proposition – the "principle of utility", as Bentham (1789/1838) called it – which does all the justificatory work. This principle says, roughly, that an act is right if and only if it maximizes total happiness.[6] Utilitarians claim that all moral questions are answered in direct or indirect reference to that first principle. Accordingly, the practical challenges of human enhancement would have to be addressed as follows. First, we agree on the criterion for rightness and wrongness. That is, we agree that an act is right if and only if it maximizes the sum total of happiness in the universe. Then, we investigate the issue at hand. We examine the choice options – i.e. whether to enhance or not to enhance a given human being in a given case – and project their consequences for overall happiness. Finally, we simply choose the option that fares better in that comparison. Case closed.

There is an alternative to Cartesian rationalism that regards our lifeworld and the normativity it is laced with as the basis of ethical theorizing. It recognizes that our practice is interwoven with norms that regulate the way we interact in our lifeworld and views ethics as an attempt to explicate that practice as well as to systematize and revise it. Austin's (1963) speech act theory may be seen as an attempt to do the former, while Rawls's (1951 and 1971/1999) "reflective equilibrium" methodology is an attempt to do the latter. The interworking of those two bodies of theory can be illustrated as follows. Take, e.g., the speech act of promising. Suppose we observe that A promises B to Φ. In that situation, we observe an empirical fact, *viz.* that A promises B to Φ. As

---

[6] For a more comprehensive statement and systematic critique of that doctrine, see Mukerji (2013).

members of our language community we understand, however, that by giving B the promise to Φ, A *ipso facto* incurs an *obligation* to Φ. In addition, therefore, we make a normative observation, *viz.* that A incurs an obligation to Φ. To that extent, then, A has a reason to Φ, which means, in all likelihood, that she should, in fact, Φ. As far as explicating the normative content of our lifeworldly institution of promising, this seems to be what we should say. Now where does reflective equilibrium come in? To see this, it is important to recognize that promises are but one aspect of our lifeworldly practice. We do believe that we should keep our promises. But we do not believe that our moral obligations are only a matter of living up to promissory obligations. There are other moral considerations that are important.[7] This is easy to see. Suppose, e.g., that by breaking her promise, A can prevent some morally horrendous event. In that case, we would certainly judge that she should break her promise. At some stage, then, the question arises how our moral judgements fit together into a coherent whole. This is where the idea of reflective equilibrium comes in. We seek reflective equilibrium when we attempt to systematize our moral judgements, that is, when we attempt to work them into a dense network of logically connected claims that make up a moral theory. Systematicity, however, is not the only guiding idea. On occasion, it may turn out that our moral judgements are inconsistent. When they are, we need to revise and reformulate at least some of them in order to make them consistent. We do that according to the subjective degree of confidence that we invest them with. That is, we keep convictions that we are rather sure about and drop conflicting judgements that seem to us less certain (Nida-Rümelin 2009).

Having introduced the two views of the edifice of moral theory, we may ask which is more appropriate. Given the scope of this essay it is, of course, impossible to give a comprehensive answer to that question.[8] For this reason, we shall confine ourselves to one brief remark, *viz.* that the rationalist view combines two rather incredible assertions. On the one hand, it holds that at least one view is absolutely certain, *viz.* the first principle that the whole of moral theory is built upon. On the other hand, there is radical doubt. Any moral judgement that collides with the first principle has to be given up. These two ideas seem maddeningly absurd. The notion that certain ideas cannot be given up *under any circumstances* appears to be plainly unreasonable. In fact,

---

[7] Certain moral considerations can override promissory obligations. Some examples can be found in Mukerji (2014).

[8] See, however, Nida-Rümelin (2009) for a thorough treatment.

this has long been recognized in the philosophy of science (see Nida-Rümelin 2009). It seems equally preposterous to suggest that we should be radically doubtful about all the moral beliefs that we currently entertain. The second view acknowledges both of these points. It is fallibilist insofar as it recognizes that any moral judgement may, in principle, be doubted. But it dispenses with radical doubt. For it acknowledges, as Wittgenstein has noted, that doubt can never be radical, but is intelligible only against the background of beliefs that are not doubted (Wittgenstein 1969). To that extent, then, our approach may be called "pragmatist".[9]

This said, let us point out that our broadly pragmatist view contains an endorsement of what Allen Buchanan calls the "Balancing Approach". This approach suggests that the right way to go about discussing issues of human enhancement (as well as any other issues) is to "look both at the considerations in favor of enhancement and those against and to strive for the judgment that reflects a proper appreciation of both." (Buchanan 2011, 59) In other words, it suggests that any radical stance either for or against enhancement should be rejected and that the question about the permissibility of human enhancement and its ethical limits is an open question that can be answered only if all the reasons – pro and con – are duly considered and appropriately weighed on a case-by-case basis.[10]

3. Human Enhancement – Pro and Con

Up until now, we have given two reasons why sweeping generalizations about human enhancement seem dubious. Firstly, it is hard to see how such a view might be supported argumentatively, given that a clear definition of enhance-ment is not available. Secondly, the pragmatist conception of moral justifi-cation that we favour supports the Balancing Approach, which suggests that we

---

[9] Note that calling a philosophical view pragmatist may mean different things. As Hookway (2013) points out, a philosophical view subscribes to "pragmatism in the narrow sense" if it accepts what C. S. Peirce's "pragmatist maxim". In contrast, it subscribes to "pragmatism in the wider sense" is it rejects the Cartesian quest for absolute certainty and accepts the fallibilist idea that "that any of our beliefs and methods could, in principle, turn out to be flawed." What we said above, makes our view pragmatist in the wider sense, while we can allow ourselves to remain agnostic about pragmatism in the narrow sense.

[10] The weighing of reasons, we should add, is a distinct capacity of human being as rational, free and responsible agents. To that extent, it cannot be algorithmized (see Nida-Rümelin, forthcoming), but demands, as Buchanan points out, a "proper appreciation" by human beings as moral agents.

should be wary of sweeping generalizations about human enhancement. But suppose, for argument's sake, that it was possible to come up with a clear and intuitively satisfactory definition. In that case, the Balancing Approach might support a general conclusion in favour of or against human enhancement. For it may turn out that all arguments line up on one side of the debate. In this section, we go through a number of normative considerations that are relevant in the context of enhancement, thereby making clear that that idea is rather doubtful.

3.1. The Human Good

The primary idea that drives supporters of human enhancement is, it seems, a concern for human welfare. Those who support enhancement mainly do so because they want human beings to live more fulfilling and better lives. As Steven Pinker points out, the advances made in the biomedical sciences, "if translated into freely undertaken treatments, could make millions of people better off and no one worse off." (Pinker 2008)

There are, of course, various views about the nature of human welfare. There is, e.g., the hedonist view, according to which a life goes well to the extent that the person living it enjoys pleasurable sensations (and is able to avoid unpleasant ones). According to the preferentist view, a life is good to the extent that the preferences of the person living it are fulfilled. And objective list theorists believe that well-being is to be seen as an index of goods. On their view, a life goes well to the extent that the person living it is healthy, emotionally intact, able to develop her capacities and so on. Since there are diverse views on the nature of human welfare, it is clear that the effects of a given enhancement technology may be judged differently depending on the theory of welfare that is used to evaluate it. Arguably though, there will be certain uses of human enhancement technologies that will increase human welfare on all plausible accounts of well-being. Increased health, e.g., is a good on all views. It is an intrinsic good on objective list accounts of well-being. The same holds on the preferentist view, at least if we can assume that every reasonable person should value health. On a hedonist view, improved health may count as an instrumental good because it is certainly one of the preconditions of a pleasurable life. There are, then, at least some welfare considerations that strongly speak in favour of certain forms of enhancement.

But welfare considerations may also be used to oppose certain types of enhancement. The extension of a human capacity does not necessarily add value to a life. This is easy to see. Suppose, e.g., that it was possible to improve the eyesight of a person up to a point where she can see at the microscopic level. This would certainly be helpful in a number of situations. But it would surely be quite a pain in many others. It would enable her to see all the microbes crawling around and the many tiny flaws in the skin of a loved one. It is thus plausible to regard the radical enhancement of our sense of vision as a doubtful "improvement" – one that is presumably not worth wanting.[11] If this example does not convince you, there are many others. The point is simply that enhancements do not necessarily increase human welfare. Whether they do depends on the empirical facts about the exact kind of enhancement that we are dealing with.

3.2. Risk

The promise of positive welfare effects seems to be the most prominent argument on the pro side of the human enhancement debate. In contrast, critics of enhancement technologies commonly emphasize their risks. The members of the President's Counsel on Bioethics, e.g., warn in their report *Beyond Therapy* (2003) that

> [t]he human body and mind, highly complex and delicately balanced as a result of eons of gradual and exacting evolution, are almost certainly at risk from any ill-considered attempt at "improvement." (President's Counsel on Bioethics, 2003, p. 287)

But even those who are generally sympathetic to the idea of human enhancement concede that the issue of risk has to be taken very seriously. This seems plausible. Even if the potential benefits of an enhancement should turn out to be enormous, it still seems inadvisable to jeopardize the comfortable level of existence that we enjoy now (at least in most parts of the Western world). This appears to hold even in the case of relatively small risks if the possible damage is significant. If we err, it seems, we had better err on the side of caution. This appears to hold, in particular, in the case of entirely new technologies that we have very little experience with. After all, even

---

[11] We owe this example to Allen Buchanan.

technologies that initially seem entirely benign may eventually turn out to involve quite significant risks.

It should be noted, of course, that considerations about the risks of enhancement do not give us a reason to oppose it categorically. They speak against a given enhancement measure only if the risk is so significant that it outweighs its expected benefits or the extent of the conceivable damage is quite large. That, however, is enough to oppose certain types of enhancements almost categorically – at least for now. It seems commonsensical, for this reason, to oppose any far-reaching interventions into the genetic makeup of humans and doubly so if these interventions can affect future generations through germ-line genetic modifications.

We may record, then, that certain considerations of risk speak against particular kinds of enhancements that are unsafe to use, at this stage anyway. But risk assessments may also be used to support a pro enhancement stance. Ingmar Persson and Julian Savulescu have done so in their recent book *Unfit for the Future* (2012). They argue that technological advances and modern-day problems, such as global warming, have tremendously increased the risks that humanity faces. E.g., in a liberal democracy the requisite knowledge to develop weapons of mass destruction is easy to come by and may, in fact, be used by a bunch of radicals in order to eradicate life on earth. In view of such immense risks, Persson and Savulescu propose to consider *moral* enhancement as a possible solution. They believe that we should try to deploy the means of biotechnology in order to make humans more moral and cooperating, thus mitigating the risks that they pose for others. There are further risk-based arguments that may be advanced in order to support certain kinds of enhancement. Nick Bostrom, e.g., homes in on health hazards. He argues that individuals who have an unenhanced genome may run risks "that can be extremely grave." To him, this is a reason that might justify genetic human enhancement, as "it would be irresponsible to risk starting someone off in life with the misfortunate of (...) an elevated susceptibility to disease." (Bostrom, 2005b, p. 212) As it turns out then, considerations regarding risk may speak for and against certain types of human enhancement.

3.3. Justice

Considerations of justice are further factors that are obviously important when it comes to issues of human enhancement. For one thing, there is the issue of

distributive justice. Critics of the use of enhancement technologies may object that these technologies will tend to favour the rich who will be able to use them to a greater extent. This will give them an even greater advantage (e.g. in the job market) and will, hence, exacerbate distributive injustices.

This argument is *prima facie* plausible, particularly when it is applied to forms of cognitive enhancement, which are likely to give those who have them a competitive advantage in the race for desirably positions in society. Note, however, that the reasoning is based on at least two assumptions. Firstly, it supposes that access to enhancements will be determined by markets which will discriminate against individuals who cannot afford them. Secondly, it assumes that enhancements will be expensive and that they will, hence, benefit only rich folks. Whether these two conditions are fulfilled depends, of course, on the specifics of the case. If there were suitable governmental programmes, enhancements could be shared out on a more egalitarian basis, such that their benefits would accrue to members of all social groups. And if the price of a given enhancement was low enough, market allocations would not be much of a problem because everybody would be able to afford their fair share. In addition, some proponents of cognitive enhancement have argued that certain psychoactive drugs may actually increase distributive justice, as they tend to work better for people at the low end of the IQ spectrum, thus mitigating a presumably unwanted source of socioeconomic inequality.

But justice not only relates to distributive socioeconomic outcomes and their inequality. Rights and liberties play a great role, too (Rawls 1971). Proponents of enhancement have used this aspect of justice to make their case. Anita Silvers, e.g., argues that

> In an era that promises enormous expansion of control over our biological processes, liberal democratic theory should prize the right of citizens to be biologically different from one another, and to diverge from species typicality – from supposed biological norms – without restrictive social penalties being imposed. (Silvers, 2008, p. 79)

And the biologist Audrey de Grey tries to garner public support for his anti-aging research by appealing to what he takes to be a "fundamental right to avoid an unnecessarily early death." (de Grey, 2005, p. 661)

Rights-based arguments, however, have not only been used by proponents of enhancement but also by critics. Francis Fukuyama, e.g., voices concerns that human enhancement might threaten the very foundation of human rights. By modifying our biology, he fears, we might "disrupt either the unity or the

continuity of human nature, and thereby the human rights that are based on it." (Fukuyama, 2002, p. 172) It seems, then, that considerations of justice may be used to support both a pro-enhancement and an anti-enhancement stance. Which considerations weigh heavier can only be decided if we pay close attention to the specifics of the individual case.

3.4. Autonomy

A forth factor that has played a great role in the enhancement debate is autonomy.[12] The basic notion behind it is that of self-government. A person may be seen as autonomous to the extent that the principles she lives by are self-imposed, such that she can conceive of herself as the *author* of her life. Autonomy is conceptually linked to other important philosophical ideas, *viz.* rationality and responsibility (Nida-Rümelin, 2001, 2005 and 2011). And it may be seen as the basis of human dignity, which is, of course, a further factor that has figured in the enhancement debate.

Jürgen Habermas has employed the notion of autonomy to make a case against genetic enhancement.[13] He argues that the prenatal genetic enhancement of a human being "changes the initial conditions for the identity formation of another person in an asymmetrical and irrevocable manner" (Habermas, 2003, p. 81) The genetic designer, argues Habermas, "makes himself the *co-author of the life of another,* he intrudes – from the interior, one could say – into the other's consciousness of her own autonomy." (ibid., p. 81) The enhanced person cannot, in other words, conceive of herself as the *sole* author of her life and thus lacks "a mental precondition for coping with the moral expectation to take (...) the sole responsibility for her own life." (ibid., p. 82).

But the idea of autonomy can also form the basis of an argument in favour of the permissibility of enhancement so long as the choice to be enhanced is made by an autonomous individual. We may believe that enhancement is risky. We may question its effects on human welfare. But if another person, who is capable of deliberating and choosing autonomously, decides that she wants to be enhanced and if this does not have any adverse effects on anybody else, then our respect for that person's autonomy commands that we leave that choice up her.

---

[12] The term "autonomy" can, of course, refer to a number of related concepts. For a disambiguation, see Christman (2011).

[13] Meanwhile, the argument has been widely discussed. For a critique, see, e.g., Bostrom (2005b).

We may conclude, then, that the various normative considerations (to do with welfare, risk, justice and autonomy) do not, in and of themselves, support a pro enhancement or an anti-enhancement stance. Discussants on both sides of the debate can draw on these factors to make their case. This, of course, does not show that every position in the debate is equally justifiable. Of course, some of the aforementioned arguments are better than others. And some ethical views about human enhancement are certainly more appropriate than others. But it seems unlikely that the most appropriate view is to think of all enhancements as permissible or to reject them all as impermissible. In fact, this would seem just as unjustifiable as the view that all medical treatments are permissible or impermissible. Obviously, a number of medical treatments are ethically justifiable. Others, however, are not. Likewise, we should expect that some enhancements are morally justifiable, while others are not. In order to find out which are which, we need to consider all morally relevant considerations and carefully weigh them on a case-by-case basis. General ethical conclusions about human enhancement are, it seems, not to be had.

Conclusion

In this essay, we have argued that radical views in favour of or against human enhancement are dubious and that we should, in effect, adopt a moderate stance. On our view, we should carefully consider the various ethical questions that arise in the context of human enhancement, approach them with an open mind and on a case-by-case basis. To make this view plausible, we showed that a clear definition of enhancement is hard to come by which, of course, makes it hard to support any sweeping conclusion about enhancement *per se*. Then, we showed that the broadly pragmatist view of the edifice of moral theory, which we take to be correct, lends further support to our thesis. Finally, we went through some important normative factors in the enhancement debate and showed that they can be used in arguments for and in arguments against enhancement. The bottom line of the reasoning that we have presented is that we should treat issues of human enhancement like we do any other ethical issue, *viz.* by carefully considering and weighing up the reasons pro and con.

REFERENCES

Allhoff, F., Lin, P., Moor, J., Weckert, J. (2009). *Ethics of Human Enhancement: 25 Questions & Answers* (Version 1.0.1) http://ethics.calpoly.edu/NSF_-report.pdf. (Accessed March 1, 2014)

Austin, J. L. (1962). *How to do things with words.* Oxford: Clarendon Press.

Bentham, J.. 1789/1838. *The Collected Works of Jeremy Bentham: Published under the Superintendence of his Exectutor John Bowring* Vol. 1. Edinburgh: Williams Tait.

Bostrom, N. (2003). The Transhumanist FAQ – A General Introduction, http://www.transhumanism.org/resources/FAQv21.pdf. (Accessed March 1, 2014)

Bostrom, N. (2005a). A History of Transhumanist Thought, *Journal of Evolution and Technology* 14(1), 1-25.

Bostrom, N. (2005b). In Defense of Posthuman Dignity, *Bioethics* 19(3), 202-214.

Bostrom, N., Roache, R. (2008). Ethical Issues in Human Enhancement, in Ryberg, J., Peterson, T., Wolf, C. (eds.). *New Waves in Applied Ethics.* Basingstoke: Palgrave Macmillan, 120-152.

Buchanan, A. (2011b). *Beyond Humanity?* Oxford: Oxford University Press.

Christman, J. (2011) "Autonomy in Moral and Political Philosophy", in Zalta, E. N. (ed.). *The Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu-/archives/spr2011/entries/autonomy-moral. (Accessed March 1, 2014)

de Grey, A. D. N. J. (2005). Life extension, human rights, and the rational refinement of repugnance, *Journal of Medical Ethics* 31: 659-663.

Descartes, R. (1641/2008). *Meditations on First Philosophy.* Translated by Michael Moriarty. Oxford: Oxford University Press.

Fukuyama, F. (2002). *Our Post Human Future.* New York: Farrar, Straus and Giroux.

Habermas, J. (2003). *The Future of Human Nature.* Cambridge: Polity Press.

Hare, R. M. (1981). *Moral Thinking – Its Levels, Method and Point.* Oxford: Oxford University Press.

Hauskeller, M. (2010). Nietzsche, the Overhuman and the Posthuman: A Reply to Stefan Sorgner, *Journal of Evolution and Technology* 21(1), 5-8.

Hookway, C. (2013) "Pragmatism", in Zalta, E. N. (ed.). *The Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/archives/win2013/entries/pragmatism. (Accessed March 1, 2014)

Kass, L. R. (2001). L'Chaim and Its Limits: Why Not Immortality?, *First Things* 113, 17-24.

Kass, L. R. (2003). Ageless Bodies, Happy Souls, *The New Atlantis* 2003(1), 9-28.

Mukerji, N. (2013). Utilitarianism, in Lütge, C. (ed.) *The Handbook of the Philosophical Foundations of Business Ethics*. Dordrecht: Springer, 297-312.

Mukerji, N. (2014). Consequentialism, Deontology, and the Morality of Promising, in Jauernig, J., Lütge, C. (eds.). *Business Ethics and Risk Management*. Dortrecht: Springer, 111-126.

Nida-Rümelin, J. (2001). *Strukturelle Rationalität*. Stuttgart: Reclam.

Nida-Rümelin, J. (2005). *Über menschliche Freiheit*. Stuttgart: Reclam.

Nida-Rümelin, J. (2009). *Philosophie und Lebensform*. Frankfurt am Main: Suhrkamp.

Nida-Rümelin, J. (2011). *Verantwortung*. Stuttgart: Reclam.

Nida-Rümelin, J. (forthcoming). Agency, Technology and Responsibility, *Politica & Società*.

Persson, I., Savulescu, J. (2012). *Unfit for the Future*. Oxford: Oxford University Press.

Pinker, S. (2008). The stupidity of dignity. *The New Republic.* http://www.tnr.com/article/the-stupidity-dignity (Accessed March 1, 2014)

President's Council on Bioethics. (2003). *Beyond Therapy*. https://bio-ethicsarchive.georgetown.edu/pcbe/reports/beyondtherapy/beyond_therapy_final_webcorrected.pdf (Accessed March 1, 2014)

Rawls, J. (1951). Outline of a decision procedure for ethics, *Philosophical* Review 60(2), 177-197.

Rawls, J. (1971/1999). *A Theory of Justice*. Cambridge, MA.: Harvard University Press.

Sandel, M. (2007). *The Case against Perfection: Ethics in the Age of Genetic Engineering.* Cambridge, MA.: Harvard University Press.

Schermer, M. (2007). The Dynamics of the Treatment-Enhancement Distinction: ADHD as a Case Study, *Philosophica* 79, 25-37.

Silvers, A. (2008). The right not to be normal as the essence of freedom, *Journal of Evolution and Technology* 18(1), 79-85.

Singer, P. (1979/1993). *Practical Ethics.* Cambridge, UK: Cambridge University Press.

Smart, J. J. C. (1973). An Outline of a System of Utilitarian Ethics. In Smart, J. J. C., Williams B. A. O. (eds.). *Utilitarianism: For and Against.* Cambridge: Cambridge University Press, 3–74.

Sorgner, S. 2009. Nietzsche, the Overhuman, and Transhumanism, *Journal of Evolution and Technology* 20(1), 29-42.

Wittgenstein, L. (1969). *On Certainty.* Edited by G.E.M. Anscombe and G. H. von Wright. Oxford: Basil Blackwell.

Transhumanism and its Genesis: The Shaping of Human Enhancement Discourse by Visions of the Future

Christopher Coenen [†]
christopher.coenen@kit.edu

ABSTRACT

Current discourse on human enhancement is strongly influenced by far-reaching, radical visions concerning the future of human corporeality and civilisation. These visions are most forcefully brought into the discussions by proponents of transhumanism, which constitutes both a worldview and a sociocultural movement that is increasingly influential in academia, industry and other sectors of society. Aiming to shed new light on our societies' current fascination with human enhancement discourse, three narratives concerning the genesis of transhumanism and the attractiveness of this worldview are presented. Such a historically interpretative approach may give rise to a new reflexive stance on current enhancement discourse.

> These bodies which now we wear belong to the lower animals; our minds have already outgrown them; already we look upon them with contempt. A time will come when Science will transform them by means which we cannot conjecture [...]. With one faith, with one desire, [men] will labour together in a Sacred Cause: the extinction of disease and sin, the perfection of genius and love, the invention of immortality, the exploration of the infinite, and the conquest of creation. (Winwood Reade, 1872)

[†] Karlsruhe Institute of Technology, Germany.

Introduction

Discourse on human enhancement is not a new phenomenon. The potentially radical impacts of new or emerging technoscience on human corporeality and civilisation only became the subject of major international discussion in the early 2000s, however. In April 2000, the computer scientist and entrepreneur Bill Joy had published an essay in which he argued that "[o]ur most powerful 21st-century technologies – robotics, genetic engineering, and nanotech – are threatening to make humans an endangered species" (Joy, 2000, n. pag.). In 2001, a remarkable workshop took place on the initiative of the U.S. National Science Foundation and the U.S. Department of Commerce. It aimed to foster a better understanding of the potentials of a number of important areas of research and development (such as the nano, bio, information, communication and neuro fields) and to promote their convergence and interplay with the overarching goal of improving human performance. The impressive list of participating individuals and institutions as well as the boldness of the claims made in the workshop proceedings (Roco and Bainbridge, 2003), in particular by the two editors, almost immediately attracted attention in various academic communities and in science and technology policy circles (cf. Kogge, 2008). Joy's essay – together with related activities pursued by a network of U.S., German and other opinion leaders in science and technology issues – had previously triggered an international debate in mass media and academia about humankind's technoscientific future (see, for example, Schirrmacher, 2001).

It soon became clear that an important role in these discussions was played by transhumanism, a worldview and sociocultural movement promoting a future in which human civilisation and corporeality have both been utterly transformed by science and technology. Transhumanism's visions of the future include not only the emergence and widespread use of enhancement technologies but also the 'uploading' of individual minds onto hardware, their quasi-telepathic interconnection and the extraterrestrial expansion of the (trans)human species.

As an organised movement, transhumanism is fairly small, yet it is not without influence, for example in academic bioethical discourse. As a worldview and broader intellectual movement, it is an important element of Western (and therefore of global) culture, for example with regard to science fiction and the sectors of popular culture influenced by this genre.

Academic and mass media discourse on human enhancement developed in the 2000s against this backdrop and in recent years appears to have been increasingly shaped by radical visions concerning the future of technoscience, human corporeality and civilisation.

In the meantime, it has also become evident that transhumanism is not only an unusual ideology and a small social movement but also a kind of worldview for significant parts of the technoscientific elite in the U.S. and elsewhere. Key players in the computer and Internet industry, for example, not only directly support transhumanists and promote their ideas but have also launched activities in their firms that have a decidedly transhumanist flavour (see, for example, Coenen et al., 2009, section 2.8.4, p. 106; McCracken and Grossman, 2013; Shanks 2013; Cadwalladr, 2014).

Many observers of and participants in discourse on human enhancement, in particular in Europe, still tend to see transhumanism's visions of the future as merely an outflow of specific developments in U.S. culture and as a kind of science fiction disguised as science. This appears to be too narrow a view of this phenomenon, however. In the early 2000s, Jürgen Habermas had already written the following:

> Bodies stuffed with prostheses to boost performance, or the intelligence of angels available on hard drives, are fantastic images. [...] Whether these speculations are manifestations of a feverish imagination or serious predictions, an expression of displaced eschatological needs or a new variety of science-fiction science, I refer to them only as examples of an instrumentalization of human nature initiating a change in the ethical self-understanding of persons who live in the mode of self-determination and responsible action. (Habermas, 2003, p. 41–42)

The question of whether transhumanist visions of the future are manifestations of a feverish imagination or serious predictions touches on an important aspect of the analysis of current discourse on human enhancement. 'Vision assessment' studies (e.g. Ferrari et al. 2012) explore how – and on the basis of what evidence and assumptions – imaginaries are influencing the creation of 'plausible futures' in academic discourse and society at large. Critics of excessively 'speculative' tendencies in ethico-political discourse on new and emerging technoscience (Nordmann 2007) have argued that transhumanism's stance towards the future is fundamentally flawed and contributes to a kind of ignorance of present challenges and options for action.

In this chapter, however, we are not concerned with these problems; instead, we would like to draw attention to questions such as that alluded to by Habermas in his remark, namely that the transhumanist visions could be an expression of displaced eschatological needs.

For this purpose, three interpretations of transhumanism will be presented, all of which aim to shed light on this worldview against broader historical and societal backgrounds (for the following sections, cf. Coenen 2013a). These interpretations, respectively, focus on a new concept of human self-assertion developed since the second half of the nineteenth century, following the Darwinian blow to human narcissism (Section 2), the dream of an empire to end all empires (Section 3), and the role that eminent natural scientists of a radically left-wing political persuasion played in shaping transhumanism as an ideology for emerging technoscience (Section 4). In offering these interpretations, which all begin by focusing on British history in the late nineteenth and early twentieth centuries, we hope to show that current discourse on human enhancement is above all a manifestation of unsolved questions which are deeply rooted in Western history of ideas about science, technology and the future (Section 5).

1. Human Self-Assertion after Darwin

In 1902, H.G. (Herbert George) Wells (1866-1946) delivered a lecture to the Royal Institution, titled The Discovery of the Future (Wells 1902). After juxtaposing two attitudes towards the past and the future, he turned his audience's attention to Positivism, arguing that we should «take into account the speculations of a certain sect and culture of people who already, before the middle of last century, had set their faces toward the future as the justifying explanation of the present» (Wells, 1902, p. 330). Referring to Auguste Comte, Wells stated that in

> the narrow and limited past he recognized men had always been like the men of to-day; in the future he could not imagine that they would be anything more than men like the men of to-day. He perceived, as we all perceive, that the old social order was breaking up, and after a richly suggestive and incomplete analysis of the forces that were breaking it up he set himself to plan a new static social order to replace it. [...] And since he could see nothing beyond man in the future, there, in that millennial fashion, Comte had to end. Since he could imagine nothing higher than man, he had to assert that humanity, and particularly the future of humanity, was the highest of all conceivable things. (Wells, 1902, p. 331)

In Wells' opinion, however, this way of looking at the human species and at the future is obsolete:

> All that was perfectly comprehensible in a thinker of the first half of the nineteenth century. But we of the early twentieth, and particularly that growing majority of us who have been born since the 'Origin of Species' was written, have no excuse for any such limited vision. [...] We perceive that man, and all the world of men, is no more than the present phase of a development so great and splendid that beside this vision epics jingle like nursery rhymes, and all the exploits of humanity shrivel to the proportion of castles in the sand. We look back through countless millions of years and see the will to live struggling out of the intertidal slime, struggling from shape to shape and from power to power, crawling and then walking confidently upon the land, struggling generation after generation to master the air, creeping down into the darkness of the deep; we see it turn upon itself in rage and hunger and reshape itself anew; we watch it draw nearer and more akin to us, expanding, elaborating itself, pursuing its relentless, inconceivable purpose, until at last it reaches us and its being beats through our brains and arteries, throbs and thunders in our battleships, roars through our cities, sings in our music, and flowers in our art. And when, from that retrospect, we turn again toward the future, surely any thought of finality, any millennial settlement of cultured persons, has vanished from our minds. This fact that man is not final is the great unmanageable, disturbing fact that arises upon us in the scientific discovery of the future, and to my mind, at any rate, the question what is to come *after* man is the most persistently fascinating and the most insoluble question in the whole world. (Wells, 1902, p. 331)

We have included this long quotation not only because it illustrates the Wellsian preoccupation with the future of human nature but also for another reason: in the passage that starts by looking back "through countless millions of years", Wells – as he did in several other parts of the lecture – evoked and addressed what can be termed a new appraisal of both the 'mathematical sublime' and the 'dynamically sublime', as defined by Immanuel Kant.

During the course of the nineteenth century, gradualist geology, Darwinism and cosmology expanded the time horizons of modernity in both directions. The distant past and the far future increasingly became subjects of inquiry and speculation. The awe-inspiring timescales and vastness of the universe created a new urgency of the mathematical sublime. As Kant (2007, p. 18) pointed out, nothing that can be an object of the senses (even using such means as telescopes or microscopes) can be described as sublime; nonetheless,

there is in our imagination a striving towards infinite progress, and in our reason a claim for absolute totality, regarded as a real idea (ibid.), and this excites in us the feeling of a supersensible faculty (ibid.). This Kantian 'dynamically sublime' appeared not only in a new light, but also against the backdrop of scientific progress during the course of the nineteenth century, since the notion of our superiority to nature even in its immensity needed to be brought up to date with respect to this backdrop.

The "kind of self-preservation, entirely different from that which can be attacked and brought into danger by external nature" is characteristic in the Kantian point of view of our stance towards the dynamically sublime, and is reaffirmed in the post-Darwinian situation. Wells and others in the late nineteenth and early twentieth centuries intended to show that "humanity in our own person" in fact remains unhumiliated by the new insights into the immenseness of timescales and vastness of space, even though the individual might have to submit to external violence (Kant 2007, p. 92). Nature is only sublime "because it elevates the imagination to a presentation of those cases in which the mind can come to feel the sublimity of its vocation even over nature" (ibid.). Human 'self-preservation' as defined by Kant is glorified by Wells and others and turned into a specific form of highly modern human self-assertion vis-à-vis the natural sublime through visions of the future which incorporate the new perspectives resulting from gradualist geology, cosmology and evolutionary theory.

One crucial element of this new concept of human self-assertion is the expectation that human corporeality will be improved, or even superseded, by a new form of corporeality. Of course, similar visions had already been developed previously, for example by Francis Bacon, the Marquis de Condorcet and William Godwin; however, Wells and others brought forward their visions of the future against the background of a new scientific appraisal of the natural sublime and with a decidedly critical stance towards older (social-)Utopian ideas about the future.

A crucial role was played in this context by Winwood Reade (1838-1875; see initial quotation above) This somewhat colourful figure, best known as a hero of the freethinker movement but also an Africa explorer in contact with Charles Darwin who influenced people as diverse as Winston Churchill, George Orwell and Wells, had published a universal history in 1872 entitled 'The Martyrdom Man' (Reade 1910). In the last two sections of this work, headed the "The future of the human race" and "The religion of reason and

love", Reade developed the blueprint for the ideological nucleus of modern transhumanism by creating a specific set of visions of and a narrative about the future of humankind. In his vision of the far future, outer space will be colonised by virtuous men endowed with new bodies created by "Science". Humanity will evolve to become a true collective, "united by the same sentiment which united the primeval clan, and which made its members think, feel, and act as one" (Reade 1910, p. 514). Reade already exhibits the following elements of modern transhumanism: the desire to overcome (by scientific means) the human body, which is seen as outmoded as compared with humanity's intellectual progress; the hope that humankind will be able to rid itself of the "stamp" of "lowly origin" in the human "bodily frame" which Darwin had mentioned and characterised as "indelible" in the final paragraph of 'The Descent of Man' (1871); the contempt for human corporeality; the quasi-religious ideological approach and opposition to traditional (Christian) religion; the extremely far future perspective; the hope for an 'invention of immortality'; and the expectation that a biologically transformed (post-)humanity will become a God-like entity ruling the universe.

It is important to note that Reade's and the Wellsian visions of the future were both based on teleological notions of progress. Reade believed, for example, that such scourges of humanity as war, famine and slavery were means used by Nature to realise historical progress but that they will, like religion, become obsolete in the future. Although he also often pointed out possible catastrophic or dystopian developments, Wells believed – as did Reade – that natural and human history are evolving towards a much greater human dominion over nature, including over human nature itself.

The early transhumanist visions of Reade and Wells were further developed by a number of important scientists like John Burdon Sanderson Haldane, Julian Huxley and, in particular, John Desmond Bernal (cf. Parrinder 1995), all of whom added to them a higher degree of technoscientific imagination. In Reade's opinion, it was "Science alone" which could "ameliorate the condition of the human race' (Reade 1910, P. 511). Wells, Haldane and Bernal departed from this point of view by more strongly emphasising the importance of political reforms, while Bernal in particular developed ideas about a technological transformation of human corporeality.

Before returning to this crucial step in the development of modern transhumanism, we would like to draw attention to Reade's two other important contributions to the emergence of transhumanism. On the one

hand, he argued that an understanding of the laws which regulate the complex phenomena of life would enable us to predict the future in the same way as we are already able to predict the movements of the planets. The above-mentioned 1902 lecture by Wells, as well as a 1929 essay by Bernal (1970) that played a crucial role in the development of modern transhumanism, followed in Reade's footsteps in this regard (cf. Parrinder 1995). On the other hand, Reade's 'The Martyrdom of Man' shows how the genesis of transhumanism has been influenced by the notion of an 'empire' and shaped by the imperialist reality of the late nineteenth and early twentieth centuries.

Before turning to this latter aspect of transhumanism's history, we would like to briefly sum up what we conclude from the analysis above: Reade and Wells put forward a new concept of human self-assertion in line with their views of scientific progress in the nineteenth century. These views were based on teleological concepts of progress concerning both natural history and social developments. As Reade wrote:

> You blessed ones who shall inherit that future age of which we can only dream; you pure and radiant beings who shall succeed us on the earth; when you turn back your eyes on us poor savages, grubbing in the ground for our daily bread, eating flesh and blood, dwelling in vile bodies which degrade us every day to a level with the beasts, tortured by pains, and by animal propensities, buried in gloomy superstitions, ignorant of Nature which yet holds us in her bonds; when you read of us in books, when you think of what we are, and compare us with yourselves, remember that it is to us you owe the foundation of your happiness and grandeur, to us who now in our libraries and laboratories and star-towers and dissecting-rooms and work-shops are preparing the materials of the human growth. (Reade, 1910, pp. 589-539)

Faced with what they saw as fundamental changes to the natural sublime driven by a new scientific outlook in the second half of the nineteenth century, Reade and Wells created awe-inspiring narratives about the past and visions of the future in which human technoscientific progress itself became endowed with features of the sublime.

2. An Empire to End all Empires

Reade wrote 'The Martyrdom of Man' after having already travelled through Africa twice. The book reflected the imperialist context of his life and activities, the author not only displaying an almost afro-centric stance but also providing

a grand narrative in which all past human endeavours and British imperialism in particular were presented as steps towards a grandiose future. All the empire-building of the past merely foreshadowed the coming (trans)human empire over nature. In 'New Atlantis' (1627; in Latin 1624), Francis Bacon had defined «the knowledge of causes, and secret motions of things; and the enlarging of the bounds of human empire, to the effecting of all things possible» (Bacon 1996, p. 480) as marking the end of all activities of the Utopian proto-technoscientific community. As Sarah Irving has pointed out with regard to Britain in early modern times, it is important to note that the term 'empire' denoted both «plenary authority or sovereignty, whether over territory or over intangible phenomena such as ideas» (Irving 2007, p. 33) and the need for us to take into account the relationships between the theory and the practice of the idea of empire in order to «reorient our thinking about the conceptual history of the British Empire». Early on in the histories of the British Empire and of science in Britain, «grand hopes for recovering man's pre-lapsarian epistemic empire» shaped discourse on the territorial empire.

In marked contrast to the usual rhetoric of "the white man's burden", Reade placed scientific progress centre-stage in his universal history, thereby influencing Churchill (Jablonsky 1991, p. 20), Wells and others. Although tensions between imperialist and anti-imperialist sentiments or ideas are evident throughout 'The Martyrdom of Man', it can be understood as an attempt to purify and perpetuate the imperial dream. Reade longed for humanity to become emancipated from the bonds of both traditional cultures and nature, and aimed to build a global, even cosmic (trans)human empire over nature. As an ideological endeavour, early transhumanism thus appears to be a project of an empire to end all (political) empires.

Similarly, Wells combined the critique of the militaristic features of imperialist societies and of other elements of the traditional social order with grandiose visions of a new and global empire based on reason, science and technocracy. There was always a tension in Wells between these visions and his liberal concerns regarding threats to individual freedom and human rights. These concerns led him, for example, to reject all 'static' Utopias and to create dystopian visions of an entirely rational, collectivist society, such as the Selenite insect society in The First Men in the Moon (1901). The Selenites have in some sense perfectly realised the Wellsian Utopia of a true empire of reason, a peaceful society giving preference to mental over physical strength yet at the same time employing brutal means to organise this very society,

subjecting for instance some of their offspring to torture and surgery in order to create slave classes that will serve the needs of the ruling intellectuals. Notwithstanding his concerns that attempts to realise his Utopia could lead to the creation of illiberal and inhumane dystopian societies, Wells forcefully promoted this very Utopia. As Dietmar Dath (2005) has pointed out, he argued in favour of socialism's transition from a romantic, static Utopianism to a bio-political technocracy, thereby extending the universalism of the British Empire to an interstellar level. In the final scene of the film 'The Shape of Things to Come' (1936), which was based on a work by Wells and whose script was written by Wells, the character that epitomises the Wellsian Utopia (and was styled after Bernal) is asked whether there will never be any age of happiness and never be any rest. In response, he declaims:

> Rest enough for the individual man: too much, too soon and we call it death. But for Man, He must go on, conquest beyond conquest. First this little planet with its winds and ways, and then all the laws of mind and matter than restrain him. Then the planets around him, and at last out across the immensity to the stars. And when he has conquered all the deeps of space and all the mysteries of time, still he will be beginning.

After the demurring remark by another character that "we're such little creatures" and "poor humanity's so fragile, so weak –little, little animals", he adds: «Little animals. If we're no more than animals we must snatch each little scrap of happiness and live and suffer and pass, mattering no more than all the other animals do or have done. It is this – or that: all the universe or nothing. […] Which shall it be?»

As Dath (2005) writes, the "civilising of the universe", imagined on the basis of a British notion of 'empire', is a vision of the future in which the emotional instincts shaping technoscientific progress are not repressed or tamed but are set free and "discharged" in a bold enterprise aimed at the complete conquest and reconstruction of the physical world.

Although Wells was not free from prejudices and sentiments characteristic of the imperialists of his times, his Utopia represents a step forward in the process by which early transhumanism became increasingly distanced from the imperialist context in which it had emerged. As George Orwell wrote about him:

> Mr. Wells […] belongs to the non-military middle class. The thunder of guns, the jingle of spurs, the catch in the throat when the old flag goes by, leave him

manifestly cold. He has an invincible hatred of the fighting, hunting, swashbuckling side of life, symbolised in all his early books by a violent propaganda against horses. [...] If one looks through nearly any book that he has written in the last forty years one finds the same idea constantly recurring: the supposed antithesis between the man of science who is working towards a planned World State and the reactionary who is trying to restore a disorderly past. [...] On the one side science, order, progress, internationalism, aeroplanes, steel, concrete, hygiene: on the other side war, nationalism, religion, monarchy, peasants, Greek professors, poets, horses. History as he sees it is a series of victories won by the scientific man over the romantic man. (Orwell, 1941, 136)

Early British transhumanism can be deemed a collectivist project aimed at perpetuating the imperial dream by purifying it. Although Wells occasionally expressed racist sentiments, one of his main lines of argument amounted to a systematic othering of the non-scientific mind. The targets of this 'othering' could be the populations of the colonies or the Jews, but at least as often were the traditional British and Western elites. Following in his footsteps, yet distancing themselves from Wells politically, the radically left-wing natural scientists Haldane and Bernal took the transhumanist attack on the old order to a new level. As communists, both not only developed a more concrete vision (in technoscientific terms) of a future in which the promises of religion have been fulfilled in this world, but also erased all racist vestiges from the transhumanist collectivist project's imperialist past. It appears to be no accident that both men spent significant amounts of their lives supporting the development of science and technology in the (former) colonies.

Unsurprisingly, we again note that transhumanism, being a quasi-Utopian ideology, is shaped by and reacts to core features of the societies in which it emerged. From this perspective, it can be deemed a techno-social imaginary which deals less with actual developments in science and technology than with the hopes and fears concerning science, technology and the future of humankind characteristic of the British, and for that matter Western, history of ideas.

3. Transhumanism as an Ideology for Technoscience

Following in the footsteps of Wells and building on a provocative and influential vision of the future developed by his friend Haldane in 'Daedalus, or Science and the Future' in 1923 (Haldane 1924), Bernal wrote 'The World,

the Flesh and the Devil. An Inquiry into the Future of the Three Enemies of the Rational Soul' in 1929 (Bernal 1970).[1]

This essay was published in the same, remarkably popular series of books as Haldane's speech and other works on science, technology and the future. Bernal's essay not only assembles earlier visions of the future by Haldane and others and develops new, technology-oriented ones; together with these older essays and works from the 1930s it also foreshadows almost all core elements of today's transhumanism (with the exception of 'cryonics', visions based on the digital revolution and, arguably, nanofuturism). In 'The World, the Flesh and the Devil', we encounter neuro-electric interfaces and the vision of a massive cyborgisation of human beings; ectogenesis (which had already been envisioned in Haldane's speech and was later popularised by Aldous Huxley's 'Brave New World'); artificial (biological) life; a quasi-immortality of individual minds in a human-machine symbiotic superstructure resembling an organism; the conquest of outer space (to which end technological solutions are described in some detail); and the universe's saturation with earth-based intelligence, an idea also familiar from today's leading transhumanist Ray Kurzweil (2005). Further core features of today's transhumanism (for example the expectation that maximum human lifespan will be significantly extended and the vision of perfect control over human emotional life) can be found in essays by other early transhumanists who were friends of or in close contact with Bernal, such as Haldane and Julian Huxley (cf. Heil, 2010).

---

[1] As has been pointed out elsewhere (Coenen 2013b; cf. Coenen, 2010), the early transhumanist works had an astounding impact on cultural discourse on science, technology and the future, one major literary reaction being the development of classic twentieth century dystopian thought. Haldane's influence on 'Brave New World' (1932) is well-known, as is the fact that this still highly influential novel was originally intended to be a direct attack on Wellsian utopianism. In a certain sense, the authors of the classical dystopian novels of the twentieth century, such as Charlotte Franken (in 'Man's World', 1926) and Yewgeny Zhamyatin and George Orwell (in 'The Road to Wigan Pier', 1937), reacted primarily to works by the proponents of early transhumanism. The same holds true for the popular Christian authors C.S. (Clive Staples) Lewis and J.R.R. (John Ronald Reuel) Tolkien. Lewis attacked Haldane in his science fiction writings, and his essay on 'The Abolition of Man' (1942) is still influential among conservative bioethicists. Tolkien's immensely popular trilogy 'The Lord of the Rings', written in the 1930s and 1940s, can in many respects be read as a critique of the transhumanist visions of Wells, Haldane, Bernal and others (Coenen, 2010; Hogan and Clarfield, 2007). Twentieth century dystopian thought and influential Christian critiques of technoscientific progress were thus deeply indebted to the early transhumanist imagination – and both continue to shape current discourse on science and technology, their ethical aspects and the future of human nature.

'Mechanical man', the product of an age-old merging of humanity – or rather, human brains – with technology which only apparently breaks with organic evolution, is actually in Bernal's view more in keeping with the true tradition of continued evolution and "the logical outcome of the type of humanity that exists at present" (Bernal, 1970, p. 42) – while 'normal man' is an evolutionary dead end.

If we look at the context of transhumanism's development in the first third of the twentieth century, we note first of all that Reade's original anti-religious impetus was intensified by its younger proponents such as the communists Bernal and Haldane. The latter not only spent considerable time engaging in polemics with Christian apologists such as C.S. Lewis and Arnold Lunn, he also appeared to have viewed the transhumanist visions of the future as his personal *ersatzreligion* throughout his life (Adams, 2000). An almost fanatic Catholic as a boy, Bernal not only alluded, perhaps mockingly, to religious ideas in the title of his 1929 essay, but also argued that, when thinking of the future, even the least religious of men all retain in their minds and cherish an idea of some transcendental, superhuman event which will bring the universe to perfection or destruction (Bernal, 1970, p. 74). While we hold the future still timidly, however, we now "perceive it for the first time, as a function of our own action" (Bernal, 1970, p. 74). As Reade already claimed, the (post)humanity of the future thinks, feels and acts as one, quasi-telepathically, yet in Bernal's essay it is a union, initially of human brains, and – at a later stage of cyborgisation – of artificially embodied minds that are interconnected by neuro-technological means.

The new 'complex minds' could "extend their perceptions and understanding and their actions far beyond those of the individual", while sense of time could be altered in dramatic ways: "the events that moved with the slowness of geological ages would be apprehended as movement, and at the same time the most rapid vibrations of the physical world could be separated" (Bernal, 1970, p. 44). Bernal describes these complex minds as 'angels' through which "the interior of the earth and the stars, the inmost cells of living things themselves, would be open to consciousness [...] and the motions of stars and living things could be directed" (ibid.). Eventually, "the heritage of the direct line of mankind – the heritage of the original life emerging on the face of the world" will disappear, "being preserved perhaps as some curious relic, while the new life which conserves none of the substance and all of the spirit of the old would take its place and continue its development" (Bernal,

1970, p. 46). Furthermore, Bernal adds, in an almost Gnostic fashion, that "consciousness itself may end or vanish in a humanity that has become completely etherealised, losing the close-knit organism, becoming masses of atoms in space communicating by radiation, and ultimately perhaps resolving itself entirely into light" (Bernal, 1970, p. 46).

The quasi-religious character of these evocations of a new technoscientific sublime is obvious, and the transhumanism of our times often follows in the footsteps of Bernal, Haldane, Reade and Wells in this regard. There is thus more than a grain of truth in current polemics against transhumanism in which religious and other critics characterise the transhumanist movement and the flights of fancy of its most radical proponents as techno-eschatological or pseudo-religious.

If we look at transhumanism's genesis in the context of a struggle between left-wing or liberal-technocratic progressives on the one side and (often right-wing) apologists of Christian religion and the old social order on the other side, however, the current discussions appear to have too narrow a perspective. The skirmishes between transhumanists and their critics today are one element of a 'culture war' over science and technology which is largely shaped by discussions in the U.S. (in particular on Darwinism) and seldom touch on fundamental questions concerning the political and social order. Although frequently overlooked today, the transhumanism of Bernal and Haldane was arguably part and parcel of a broader fight against the old social order of their times; this fight, in which both men acted as famous 'red scientists' (Werskey 2007), was not merely a 'culture war' but a conflict which shaped the twentieth century and became a global 'cold war' after the victory over fascist Germany and its allies.

When some of today's scientists resort to transhumanist ideology and visions of the future in an attempt to counter critiques by or to provoke religious conservatives, ecological activists and other ideological adversaries, however, the situation is very different. Given the new surge in religiously framed and politically influential irrationalism in recent decades, something that is not limited to the countries shaped by Islam but is also evident in the U.S., the stance taken by transhumanist scientists is psychologically understandable. The means they use, however, and indeed often their posture, are also problematic. They are now acting in a global context in which societies are strongly shaped by technoscience and in which scientists and engineers, enjoy very considerable social status. Criticisms of science and technology

aimed at taming technoscientific progress are also societally relevant, yet they clearly do not represent the view of the political establishment nor of other leading representatives of the social order – as was the case to a certain extent in the Britain of Haldane's and Bernal's time. Quite on the contrary, transhumanism is increasingly the ideology of choice among important members of the societal elites such as several leading figures of the U.S. computer and Internet industry. As has been argued elsewhere (Coenen 2013b), discourse on some fields of new and emerging technoscience and on human enhancement in particular suffers from an ideological imbalance: at one end of the spectrum of opinions, the one marked by radical criticisms of technoscience, a barrier is erected against unscientific beliefs and fundamentalist currents of thought. At the other end of the spectrum, however, the one which is strongly influenced by transhumanism, the limits to salvation ideologies and mythical thought are permeable.

There is another element of early transhumanism's quasi-religious character which should be taken into account when considering current discourse on human enhancement, namely the specific notion of progress on which the transhumanist visions of the future are based. Haldane, for example, opined in the 1920s that "there is no theoretical limit to man's material progress but the subjection to complete conscious control of every atom and every quantum of radiation in the universe" and that "there is, perhaps, no limit at all to his intellectual and spiritual progress" (Haldane, 1937, p. 144). In 'The World, the Flesh and the Devil', Bernal repeatedly, after the boldest flights of fancy, stopped short of defining an end to progress, pointing out that another, as yet unimaginable progress might be feasible. In his view, we should not stop our imaginations until they are fully exhausted: while the fulfilment of all major promises of traditional (Christian) religion is in fact already imaginable on the basis of scientific prophecy and all eschatological questions are thereby solved, progress must never end. In adopting this stance towards the future, Bernal and the other pioneers of transhumanism not only created an alternative to Christian religion and, as we will see below, touched on unresolved questions in the Western history of ideas about science, technology and the future; they also supported a broader agenda for firmly establishing technoscience in society. As Haldane pointed out in the early 1930s, the transhumanist visions of the future served specific purposes:

> Such speculations as these are very far from idle. They are eminently desirable, because man does not generally even know what he wants, much less how to get

it. A discussion of possibilities will have two effects. It will enable people to
come to some opinions as to the possible goal of human evolution [...]. And it
will focus attention on the necessity for more knowledge before we can even
suggest means of attaining that goal. Pictures of the future are myths, but myths
have a very real influence in the present. [...] Our greatest living mythologist,
Wells, is certainly influencing the history of the future, though probably in
ways which he does not suspect. The time will probably come when men in
general accept the future evolution of their species as a probable fact, just as to-
day they accept the idea of social and political progress. We cannot say how this
idea will affect them. We can be sure that if it is accepted it will have vast effects.
It is the business of mythologists to-day to present that idea. (Haldane, 1937,
pp. 98-99)

As an ideology for emerging technoscience, early transhumanism
consciously aimed to weaken traditional religion's grip on elites and, more
broadly, on society, and to fascinate the public by offering a vision of the future
in which humans and their societies have been utterly transformed by means of
science and technology. For these purposes, the pioneers of transhumanism
also used certain radical postures and elements of cultural history. Reade had
already styled himself as, and in some sense had indeed been, an adventurer at
the fringes of society and a martyr for a great cause. In 'Daedalus', Haldane
wrote:

The conservative has but little to fear from the man whose reason is the servant
of his passions, but let him beware of him in whom reason has become the
greatest and most terrible of the passions. These are the wreckers of outworn
empires and civilisations, doubters, disintegrators, decides. In the past they
have been, in general, men like Voltaire, Bentham, Thales, Marx, and very
possibly the divine Julius, but I think that Darwin furnishes an example of the
same relentlessness of reason in the field of science. I suspect that as it
becomes clear that at present reason not only has a freer play in science than
elsewhere, but can produce as great effects on the world through science as
through politics, philosophy, or literature, there will be more Darwins. Such
men are interested primarily in truth as such, but they can hardly be quite
uninterested in what will happen when they throw down their dragon's teeth
into the world. I do not say that biologists as a general rule try to imagine in any
detail the future applications of their science. The central problems of life for
them may be the relationship between the echinoderms and the brachiopods,
and the attempt to live on their salaries. They do not see themselves as sinister
and revolutionary figures. They have no time to dream. But I suspect that more
of them dream than would care to confess it. (Haldane, 1924, p. 83-84)

Although he tempered his technocratic fervour in his more Marxist writings after the 1920s, Bernal adopted a similar posture in 'The World, the Flesh and the Devil'.[2] Also in line with the Wellsian ideas developed in 'The Discovery of Future', a restless, forward-looking and future-oriented technoscientific elite is juxtaposed with the conservative parts of humanity among which the followers of old-fashioned social Utopianism are also counted. In the future, the latter will happily live in a Utopian global society on earth characterised by Bernal as a "human zoo", and will be secretly controlled by an extra-terrestrial technoscientific elite that has been turned into cyborgs. In his essay, Bernal also took advantage of the Gothic fearful delight with the monstrous and long-standing fascination with mechanical beings. In his view, the 'mechanical man' he envisioned as the man of the future must "appear to those who have not contemplated him before as a strange, monstrous and inhuman creature" (Bernal, 1970, p. 73).

From the start, transhumanism exhibited a radical posture against not only the old social order but also against purportedly old-fashioned social Utopianism. Again, there is thus more than a grain of truth in the many criticisms of transhumanism which focus on what the critics see as its ignorance concerning the real obstacles to human progress, such as global injustice. Today, in a time when transhumanism has become the favourite ideology of parts of the digital upper class, such criticisms suggest themselves, but the fact that a certain imbalance between social and technoscientific progress was already evident in the early transhumanist visions propelled by radical leftists such as Bernal and Haldane shows that the transhumanist vision of human progress is in fact not limited to this world but is above all an expression of metaphysical concerns and eschatological needs.

When looking at the history of transhumanism in the first third of the twentieth century, it appears as an ideology for emerging technoscience, creating a new kind of sublime in an attempt to combat social obstacles to technoscientific progress as perceived by the early transhumanists. Bernal, Haldane, Wells, Julian Huxley and other pioneers of transhumanism were heavily involved in a wide variety of activities aimed at improving the societal role of applied science, the funding and organisation of science, and its public

---

[2] Neither Bernal nor Haldane ever abandoned their transhumanist visions, still adhering to them even once they had become globally prominent 'red scientists' – the best-remembered of their manifold roles. Even in the late 1950s and early 1960s, they were still publishing radically transhumanist visions of the future (Bernal 1958; Haldane 1963).

communication. Even some of their boldest visions of the future breathe the spirit of emerging technoscience. Bernal's species (techno-)brain is, for example, structured internally in a hierarchical manner, since "to some minds might be delegated the task of ensuring the proper functioning of the others, some might specialise in sense reception and so on" (Bernal, 1970, p. 44). Thus a "hierarchy of minds" would evolve, Bernal argued, in a fashion similar to Wellsian technoscientific elitism and foreshadowing the 'cybermind' visions of our times, driven forward by the likes of Hans Moravec and Marvin Minsky.

William Bainbridge, a science manager, sociologist and important transhumanist of our times, wrote that "[i]n the distant future, we may learn to conceptualize our biological lives on Earth as extended childhoods preparing us for the real life that follows in cyberspace" and that "the transition from flesh to data will not be so much metamorphosis as liberation" (Bainbridge, 2004, p. 119). In his vision, we "will travel across immensity" as "information contained in a star-spanning database", creating "new bodies along the way to dwell in every possible environment, and have adventures of the spirit throughout the universe" (ibid.). We should, Bainbridge wrote, "no more lament the loss of the bodies that we leave behind than an eagle hatchling laments the shattered fragments of its egg when it first takes wing" (ibid.).

While such remnants of the heroic youth of technoscience may appear somewhat out of place in our times, their increasing relevance in current discourse on science, technology and the future shows not only that the technoscientific sublime created since the 1870s is still with us, but that it remains at the heart of technoscience. Apparently, today's global players such as Google still follow an agenda which was developed in Britain in the heyday of imperialism and after the Great War as a reaction to a perceived crisis of progressive thinking and as a contribution to the establishment of technoscience in society. Notwithstanding its focus on individual choices, the ideological foundations of current discourse on human enhancement are collectivistic. Due to their ultimately eschatological orientation, the transhumanist visions of the future could and can be elements of politically quite different projects, such as British imperialism, communism and 'digital capitalism'. Nevertheless, current transhumanism, as an ideology for technoscience, always expresses the belief in a grand narrative about the future of humankind and thereby also shapes discourse on human enhancement in a way that allows science and technology to appear as the means of individual and collective salvation.

Conclusion

It is important to note that the grand transhumanist narrative about science, technology and the future of human nature does not form the ideological basis for all social practices and cultural movements relevant to current discourse on more radical visions of human enhancement. The rather old-fashioned modern-progressive ideology of transhumanism is but one element of a variety of developments in which human corporeality is re-defined and re-designed with a view to finding new ways in which to merge humans and technologies. Not all of the new cyborgs (see, for example, Lanxon, 2012) who technologically often follow in the footsteps of either the researcher Kevin Warwick or the performance artist Stelarc are transhumanists. Ideologically, some of them are inspired for example by Donna Haraway's cyborg feminism (cf. Haraway 1991) or other intellectual traditions that are critical of "classically" modern beliefs in progress. Current cyborgism may lead to a diversity of body modifications, not all of them compatible with the grand transhumanist narrative about the future. Some of the new cyborgs and certain voices in current discourse on human enhancement criticise or do not place the emphasis on attempts to improve human performance along the usual transhumanist lines. Attempts to extend human bodily faculties by means of implants and other devices do not always seek to improve individual competitiveness in capitalist society, and many users or promoters of cyborg technologies are not aware of or indeed reject the transhumanist narrative which regards the use of these technologies as one of the first steps towards a transhuman species.

Transhumanism – and therefore large parts of discourse on human enhancement – can be deemed both a symptom of and an attempt to end the overstraining of the rationality of the modern idea of progress diagnosed by Hans Blumenberg (1983). This overstraining by eschatological anxieties and hopes led to the "transformation of progress into a faith encompassing the future" (Blumenberg 1983, p. 49). Modern self-assertion is not only based on the progress that has been made concerning the means of self-preservation for individuals and the species at large. Against the background of the Western history of religious ideas, modern self-assertion has also always implied adopting a new stance towards the eschatological questions inherited from the Christian past. Given that some members of the technoscientific and intellectual elites are virtually obsessed with visions of radically new means of

self-preservation (such as 'cybernetic immortality') nowadays, in an era in which the scourges of humanity that Reade had believed to have been almost overcome are still with us, we may be well-advised to rethink the way in which modern self-assertion is entangled with both eschatological anxieties or hopes and our understandings of human self-preservation. As proponents of the Frankfurt School have persuasively argued, the fearful obsession with self-preservation is anachronistic and prevents us from realising the truly progressive potential that is latent in modern society. In light of the strange fact that the grand transhumanist narrative about the future has fascinated and continues to fascinate representatives of a wide variety of political persuasions, discourse on human enhancement should be redirected. The technoscientific sublime, which was created by the pioneers of transhumanism and appears to be increasingly influential in our times, is the fearful obsession with human self-preservation writ large. We may be able to break its spell by more fundamentally questioning its focus on abilities than was done in discourse on human enhancement. Gregor Wolbring has pointed out in numerous publications (see, for example, Wolbring 2007) that the 'transhumanisation of ableism', i.e. the establishment of the 'enhanced' human body as the new norm, is merely the most recent manifestation of a more fundamental problem shaping our societies, namely our obsession with abilities. If we could learn not to judge individuals by their abilities (and not to reduce them conceptually to these abilities), we might also be able to develop a vision of our common future that differs qualitatively from the transhumanist cosmic prophecies, renouncing all manner of empire-building and gloomy evocations of the technoscientific sublime.

REFERENCES

Adams, M.B. (2000). Last Judgment: The Visionary Biology of J.B.S. Haldane. *Journal of the History of Biology*, *33*/3, 457-491.

Bacon, F. (1996). New Atlantis (orig. in Latin 1624). In Vickers, B. (Ed.): *Francis Bacon: A Critical Edition of the Major Works*. Oxford: Oxford University Press, 457-489.

Bainbridge, W. (2004). Progress toward Cyberimmortality. In Immortality Institute (Ed.): *The Scientific Conquest of Death*. Buenos Aires: Libros en Red., pp. 107-122.

Bernal, J.D. (1958). *World Without War*. London: Routledge & Kegan Paul.

Bernal, J.D. (1970). *The World, The Flesh and The Devil. An Inquiry into the Future of the Three Enemies of the Rational Soul* (orig. 1929). London: Jonathan Cape.

Blumenberg, H. (1983). *The Legitimacy of the Modern Age* (orig. in German 1966). Cambridge/MA, London: MIT Press.

Cadwalladr, C. (2014). Are the robots about to rise? Google's new director of engineering thinks so... *The Observer*, (22 February 2014) [online: http://www.theguardian.com/technology/2014/feb/22/robots-google-ray-kurzweil-terminator-singularity-artificial-intelligence; last checked 12 March 2014]

Coenen, C., Schuijff, M., Smits, M., Klaassen, P., Hennen, L., Rader, M., Wolbring, G. (2009). *Human enhancement* (IP/A/STOA/FWC/2005-28/SC32 & 39). Brussels: European Parliament. [online: http://www.itas.kit.edu/pub-/v/2009/coua09a.pdf; last checked 12 March 2014]

Coenen, C. (2010). Zum mythischen Kontext der Debatte über Human Enhancement. In Coenen, C., Gammel, S., Heil, R., Woyke, A. (Eds.): *Die Debatte über „Human Enhancement". Historische, philosophische und ethische Aspekte der technologischen Verbesserung des Menschen*. Bielefeld: transcript 2010, 63-90.

Coenen, C. (2012). Cosa verrà dopo l'uomo? Osservazioni sulla storia del postumano, 159-174. In L. Grion, L. (Ed.): *La sfida postumanista. Colloqui sul significato della tecnica*. Bologna: Il Mulino 2012, 149-174.

Coenen, C. (2013a). Nachdarwinsche Visionen einer technischen Transformation der Menschheit. In Ebert, U., Riha, O. Zerling, L. (Eds.): *Der Mensch der Zukunft – Hintergründe, Ziele und Probleme des Human Enhancement*. Stuttgart, Leipzig (Abhandlungen der Sächsischen Akademie der Wissenschaften 82/3): Hirzel, 9-36.

Coenen, C. (2013b). Human enhancement. In Fischer, R., Boer, T. (Eds.): *Human enhancement: Scientific, ethical and theological aspects from a European perspective*. Strasbourg: Church and Society Commission of CEC 2013, 57-79. [online: http://www.ceceurope.org/fileadmin/filer/csc/Ethics_-Biotechnology/Human_Enhancement/CEC-Bookonline.pdf; Accessed 12 March 2014]

Dath, D. (2005). H.G. Wells: Imperialist der Vernunft. *Frankfurter Allgemeine Zeitung* 146 (28 June 2005), 37.

Ferrari, A., Coenen, C., Grunwald, A. (2012). Visions and ethics in current discourse on human enhancement. *NanoEthics*, *6*(3), 215-229.

Habermas. J. (2003): *The Future of Human Nature* (orig. in German 2001). Cambridge: Polity Press.

Haldane, J.B.S. (1924). *Daedalus; or, Science and the future: a paper read to the Heretics*, Cambridge, on February 4th, 1923. New York: E.P. Dutton & Company.

Haldane, J.B.S. (1937): *The Inequality of Man* (orig. 1932). Harmondsworth: Pelican Books.

Haldane, J.B.S. (1963). Biological Possibilities for the Human Species in the Next Ten Thousand Years. In Wolstenholme, G. (Ed.): *Man and His Future*. Boston, Toronto: Little, Brown and Company, 337-361.

Donna Haraway (1991). A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century (orig. 1985). In Haraway, D.: *Simians, Cyborgs and Women: The Reinvention of Nature*. New York: Routledge, 149-181.

Heil, R. (2010). Human Enhancement – Eine Motivsuche bei J.D. Bernal, J.B.S. Haldane und J. Huxley. In Coenen, C., Gammel, S., Heil, R., Woyke, A. (Eds.): *Die Debatte über „Human Enhancement". Historische, philosophische und ethische Aspekte der technologischen Verbesserung des Menschen*. Bielefeld: transcript 2010, 41-62.

Hogan, D., Clarfield, M. (2007). Venerable or vulnerable: ageing and old age in JRR Tolkien's The lord of the rings. *Medical Humanities*, *33*, 5-10.

Irving, S. (2007). An Empire restored: America and the Royal Society of London in the Restoration. In Armstrong, C., Fagge, R. and Lockley, T.J. (Eds.): *America in the British Imagination*. Newcastle: Cambridge Scholars Publishing, 27-47.

Joy, B. (2000). Why the future doesn't need us. *Wired, 8.04* (April). [retrieved from http://www.wired.com/wired/archive/8.04/joy.html; last checked 12 March 2014]

Jablonsky, D. (1991): *Churchill, The Great Game and Total War*. London, New York: Routledge.

Kant, Immanuel (2007). *Critique of Judgment* (orig. in German 1790). Oxford: Oxford University Press.

Kogge, W. (2008). Technologie des 21. Jahrhunderts. Perspektiven der Technikphilosophie. *Deutsche Zeitschrift für Philosophie*, 56(6), 935-956.

Lanxon, N. (2012). Practical transhumanism: five living cyborgs. *Wired* (4 September 2012). [online: http://www.wired.co.uk/news/archive/2012-09/04/cyborgs; Accessed 12 March 2014]

McCracken, H., Grossman, L. (2013): Google vs. Death. *Time* (30 September 2013; title story).

Nordmann, A. (2007). If and Then: A Critique of Speculative NanoEthics. *NanoEthics, 1*(1), 31-46.

Orwell, G. (1941). Wells, Hitler, and the World State. *Horizon, IV*(20), 133-139.

Parrinder, P. (1995). *Shadows of the Future. H.G. Wells, Science Fiction and Prophecy*. Liverpool: Liverpool University Press.

Reade, W. (1910). *The Martyrdom of Man* (orig. 1872; 18th edition with an introduction by F. Legge). London: Kegan Paul, Trench, Trübner & Co.

Roco, M., Bainbridge, W. (Eds.) (2003): *Converging Technologies for Improving Human Performance*. Dordrecht: Springer.

Schirrmacher, F. (Ed.) (2001). *Die Darwin AG*. Köln: Kiepenheuer & Witsch.

Shanks, P. (2013). Google vs. Death? Really? *Huffington Post* (10 October 2013) [online: http://www.huffingtonpost.com/pete-shanks/google-vs-death-really_b_4020540.html; last checked 12 Mrach 2014]

Wells, H.G. (1902). The Discovery of the Future. A Discourse Delivered to the Royal Institution on January 24, 1902. *Nature, 65*(1684), 326-331.

Werskey, G. (2007). The Visible College Revisited: Second Opinions on the Red Scientists of the 1930s. *Minerva*, 45/3, 305-319.

Wolbring, G. (2007). Ableism, Enhancement Medicine and the Techno-Poor Disabled In Unnatural. In Healey, P. and Rayner, S. (Eds.): *Selection: The Challenges of Engineering Tomorrow's People.* London: Earthscan, 196–209.

Human Enhancement and the Post-Human; the Converging and Diverging Pathways of Human, Hybrid and Artificial Anthropoids

Barbara Henry [†]
henry@sssup.it

ABSTRACT

The expression "human enhancement" could be placed in the ontological, cognitive, and symbolic dimension in which we conceive and experience the faculty, that is constitutive of human beings, of giving name and thus consistence to things, relations and phenomena in general. It is necessary to point out that this symbolic dimension of emerging technologies has been obstinately and jealously anthropocentric, at least in the modern Western world. In this contribution, I aim to develop a philosophical account of post-human enhancement that allows us to conceive a future society of humanoids – humans, hybrids, artificial beings – who are free and equal. This expression – "post-human enhancement" – is to be understood as referring to symbols and phenomena different from those associated with "trans-human". Post-human is to be interpreted here as material, not anthropocentric but rather interspeciesist, osmotic and relational, horizon of effective sharing of experiences, dangers and challenges. In contrast, trans-human is meant to refer to the transcending of humans into the pure ether of an 'ideal', immaterial network made up only of software, and lacking of relations with any material beings in the ecosystem or cosmos. On my account, reframing the debate about human enhancement means to guarantee widest possible conditions of non-hegemonic or expansive conscious contextuality of legislative and decisional systems. I focus rather on the social circumstance whereby we see ourselves as subjects that already co-inhabit multiform social identities, in changeable and hybrid bodies and identitary images, in potential or latent conditions of moral and political asymmetry. These conditions, I hold, are therefore to be preventively identified and neutralized.

[†] Scuola Superiore Sant'Anna, Italy.

Introduction

The "strengthening and improvement of the human being by means of targeted technological, specific, repeatable and measurable interventions, in specific areas (for example, body, behaviour, personality, intelligence), and directed at reaching stages of efficiency or adequacy or excellence, or completeness or beauty greater than the normal" can be defined as *Human Enhancement*. The noun alone indicates the equipment suited to increasing or improving the quality of a performance, of a value, or of a status. Furthermore, under certain, yet not entirely clear, circumstances, this can be equivalent to *Optimierung*, expressed with a term borrowed from the second natural western language, other than *Anglo-American*, in which there is propagation of scientific formulations, technological applications and the most sophisticated reflections concerning the field of investigation in question. Having said that, not all the procedures of optimization (*Optimierung*) can in fact be referred to as *Enhancement* (Straub, in Sieben, Sabisch-Fechtelpelter, Straub, 2012). For example those innovative and accelerating remedies aimed at the mere restoration of a condition of health with respect to the 'normal' standard. Therapy, no matter how sophisticated, cannot generally be considered enhancement, except for all those cases in which we produce, in certain conditions of connectivity between the human and technical sides, a qualitative surplus capable of endowing the 'patient' with superhuman skills, for example due to the *holistic and recursive* link existing between some kinds of prosthesis and the host organism (Coenen, 2012). In this case, the borders between therapy, *Optimierung* and enhancement tend to vanish. It is therefore only true on paper that the aim of *overcoming/going beyond*, of *ulteriority*, with all the distinctions that we shall see, and not that of reintegrating a condition of health – or lost normality, is the primary and exclusive characteristic of human enhancement.

As should clearly be understandable, in order to face the theme, from any point of view and for any pertinent context, it is necessary as a first step to presuppose a standard. There is need for a criterion, if not of normality, at least of a generalised statistical shared agreement of certain qualities or characteristics, taken for granted as indisputable or self-evident and therefore binding as comparative criteria. The aspect of regulation or legislation, prescriptive by extension, here alludes not so much to the obligation of following the rules of a norm of behaviour, which cannot be excluded, as we

shall see from the range of debates on 'human enhancement', but to that typical of the hypothetical or technical imperative. According to this type of prescription, given an x, at this point in time we are dealing with the basic starting point and point of comparison, and if we want to reach the level higher than x, termed y, we must do this or that, with these methods and these means, in relation to that specific subject, according to this timing, and so on.

No one usually asks the question of whether 'the starting x' can be perceived and identified, in the average number of cases, according to a logic, a grammar, a semantic code that is different from those that make them what they are. Things are postponed and with implicit automatism, delegated to 'normal science', to the acquisitions and knowledge that make up the background and framework of our daily certainties. This statement of the existence of a minimum benchmark that is true for a good majority of human beings imposes a connoted and committed declaration which, in turn, imposes a change in the coordinates and direction of our thinking.

1. Occidentalism and Human Enhancement. For a Change of Perspective

The «power to name beings» is that which has so far enabled us human beings to self-define ourselves as such, therefore to be positioned cognitively and strategically in the border zone between nature and the artificial. The (ontological, cognitive and symbolic) dimension in which, till now, we have conceived and experienced this constitutive faculty of giving name and consistence to things, relations, phenomena in general, has been moreover obstinately and jealously anthropocentric, at least in the modern Western world. This is also the contextual location of theoretical-political value in which, in the *most recent* centuries this poietic-classificatory vision has been used by the holders of symbolic power with unopposed supremacy with respect to other regions of the world. These in turn have been orientated, despite the potestative and definatory irrelevance lasting for some centuries and not to the same extent everywhere, towards cosmocentric, holistic and osmotic visions, as regards the levels and configurations of being. Those that we are alluding to are more or less the same societies that were hastily dismissed, with due differentiations up until the 1960s, by the codified jargon in the sociology of modernization and development manuals, as primitive, atavic, traditional and backward. This was at least until, in the more fortunate cases, they had

undergone a process of progress, according to prefixed indicators, along the lines of a (moreover, taken for granted) united pathway with the West.

The statement "we are in an era of Westernism and no longer of Orientalism" expresses on the other hand the hope (not conviction) that the time is ripe for the power-knowledge duo to be re-elaborated, despite permanent restraints and resistances, within an 'intercultural context'. This is so that the power and the pragmatic and disciplinary knowledge, including studies of political theory and international relations (Rivera, 2005, Labanca, in Cavallarin, Henry, 2012) assume rigorously more inclusive characteristics not only with respect to geo-political and cultural contexts until now directed at 'our' eyes by Orientalistic definatory opacities, *but also with respect to 'non' human and 'non' organic spheres of material existence*. These former spheres have for centuries been part of the imaginary and conditions of life and material and symbolic exchange in those very same regions of the world (Henry, 2011a, 2013). Paradigmatic examples: the circumstance of inorganic matter, in all its configurations, and automatons are very much at home together in modern Japanese society. These are the derivates of a vision and image of the world that is infraspecist and holistic, internally differentiated and internally communicating, neither anthropocentric nor Cartesian, which alone lets us understand the solution given by Japan to the challenges presented by the automation of social and working processes. The aspiration to see in the near future androids (and not only the non-mimetic robots like Asimo)[1] strolling along the streets together with us, in the role of interlocutors and not only servants within a society of relational beings much wider than the current one, is neither the dream of a visionary, nor that of a fanciful screenwriter of *anime* or *manga* (Taganishi, 2008). On the contrary, it is the strategic aim and "social" *raison d'être* of FuRo – Future Robotics Technology Center, branch of the Chiba Institute of Technology, one of the most prestigious and competitive centres of Japanese robotics research. The future has a very ancient nucleus, which is nevertheless a very pulsating one.

That is to say, in other modernizations that cannot be assimilated in a single model, and moreover hard to identify in the West, there immediately stand out particular bonds between *enhancement and a certain accepted meaning of*

---

[1] This is the name, given by his creators in honour of the father of robotic science fiction Isaac Asimov, to the small 'service' robot, similar to an astronaut – a child who welcomes with courteous phrases and bows, recognising and showing the way to the guests who arrive at the research institute in which he was created.

post-humanism, to be taken in the holistic, metamorphic and osmotic definition, not anthropocentric but rather interspecist, and anti-dualistic, as mentioned above. On the one hand, in these narratives of bonds and alliances among species and the dimensions of reality there is clear evidence of the characteristics of copresence, contiguity, transitivity, co-belonging of the different levels and forms of materiality and life. On the other hand, the ontological and axiological dualisms that are opposed in those images of the world are manifold. It is the dichotomy between both rational-spiritual dimension and material dimension, and also between immanent dimension and transcendent dimension, and also between humanity and other organic and inorganic forms of existence. In fact, these visions of reality and pragmatic images of the world, such as Shinto, Taoism and other sophisticated versions with much more ancient animist roots, do not reveal 'a night in which all the cows are black' (Henry, 2011b). On the contrary, they presuppose and legitimize systems of relations that are symbolic-material, stratified and structured axiological and potestative, according to sophisticated and pondered taxonomies, put to the test for centuries by social repercussions within the respective collective contexts.

Not even such an accepted meaning and modality of realization of human enhancement is immune to shadows or risks of irenicism, of fatuous and pernicious 'technophile' optimism. However, it is NOT legitimate for it to be identified *tout court* with a supermanistic and anthropocentric vision, bound by relations with the ecosystem, non-human species, matter and the cosmos, of technological enhancement applied to the human dimension.

2. Post-human Enhancement versus Trans-human Enhancement

The poietic-cognitive and lexicographical dimension, that is of thought, of language and codes, of experience and action, in which the basic nomenclature is situated, widened in the intercultural and interreligious sense described above, defines – *including the human, and the organic, non-organic and artificial non-human,* the *post*-human horizon.

This expression – post-human horizon – is to be understood as referring to symbols and phenomena different from those associated with trans-humanism, in which, for that matter, the adjective 'post-human' is often used in the sense of transcending what we are now. The trans-humanism alluded to here is to be understood in the accepted meaning, reconstructable on the basis of texts and

hyper-texts, of an ideology directed towards the passing/abolition of the status of 'being human' in terms of finite and incarnate living beings, of which the myth/prevision of *mind uploading* (Moravec, 1988) is only the first step. The trans-human condition is coherently intended by its proponents as "phase of transition between or animal ancestry and our post-human future" (www.extropy.org/principles.htm). The objective of overcoming – meant as a technological going beyond carried out by the human being towards a further stage of evolution – is indicated by proposers of the doctrine as being pursuable by means of the systematic struggle against the limits of our condition, and through cybernetic strategies which are configured in radical cases as *dematerializing* procedures.

However, also in the version that is more 'moderate' and closest to the cult, mediated by the productive imperatives of *cultural capitalism*, of perpetual physical and mental youth (Esfandiary, 1973; www.extropy.org, http://transhumanism.org), these practices, measures, policies, technologies, hypothesized and/or designed, always address the constitutive imperfection of our species: the finitude, the helplessness when faced with the most serious cases and adversities, the conclusive and irreversible decay and caducity of our body and faculties (Caronia, 2008, Coenen, 2009; Woyke, 2010 in Coenen, Gammel Heil, Woyke). Mortality is the enemy, in particular the burden (not the mind) is seen as an inscribed seal, object to be bemoaned, and entry point of Θανατος.

One of the most emblematic myths which has been circulating for almost a century as narration anticipating the specific most radical trans-human tendency (dematerializing) is not so much extreme physical improvement (with every possible technological means and through original contaminations between 'organic' and 'mechanical'), but rather the identitary fusion of the thinking individuals of the Web and the overcoming of their corporeity[2] Anticipations of this myth of transcendence (annihilation) of the body in pure mental and cerebral functionality can be found in the literature of the 1920s and '30s of the XXth century (Bernal, 1929, pp. 29, 43).

In this case we would have an overcoming that is understood and pursued as definitive annulment of what 'was' human, due to the engineered manufacture

---

[2] Valid examples due to their diverse versions of the global imaginary include the plot of the film *The Lawnmower Man* (not of the short story by S. King, on which it was based), the *cyberpunk* programme and it derivatives, both taken to some of its extremes (Gibson, 1984, 1987, 1988), and some episodes of the first Star Trek series.

of a humanity transcended not only in other than itself but in an alterity immeasurable with the material dimension, and capable of technologically realizing the condition that has for centuries been described as transcending dimension. In order to give contrastive examples still taken from the religious tradition of the monotheisms, the most usual symbolic reservoirs of myths of immortality for westerners, we can consider the following: according to the trans-human profession, immortality, the ultimate end, is not even reached with the transfiguration of the body (the myth of Ezekiel's prophecy, or of the Christian resurrection of the dead on the Day of Judgement), but with its cancellation, in favour of the elevation of the spirit – mental faculties (the immortal soul, the *neshamah*, the third and incorruptible form of the soul, in the Hebrew lexicon), in a rational dimension. This latter is relational, certainly, but in the sense of communication among incorporeal beings, as are the angels, or the eons, the intermediate rational beings of which Kant speaks in order to exemplify the model of quintessential *egalitarian* republic, in which all rational beings (in this case, material and immaterial) are equal, with the only exception of God. In the radical trans-human transposition the place of God, in the relation between immaterial equals with a unique unequal and supreme immaterial, is taken by the Web.

Conversely, the post-human is to be understood here as material horizon of effective sharing, that is built on the deepened knowledge and on the grounded interpretation of differences, of vulnerability, of finitude, and of the creaturality of all organic and inorganic beings, both natural and artificial. Above all, as regards the latter class of *unprecedented*, or at least unfamiliar to most, class of beings, such an objective should be pursued without confusion of category, because it can be reached only by means of a rigorous conceptual and definatory analysis.

The problem is that human enhancement can take both paths, post-human or radical trans-human.

In the first case, the overcoming of the limit is equivalent to trespassing in other territories, heterogeneous but related and not preconceptually hostile or incompatible. In the second case, the same term indicates the overcoming/annulment of the human and *ontic* condition as we knew it, the transcending of humans into the pure ether of an 'ideal' network made up only of software, and lacking relations with any material being in the ecosystem or cosmos.

For this reason it has been necessary to undergo a preliminary category distinction between post-human condition and dematerializing trans-human condition. We must now proceed by adopting a 'minimalist' and counterpoint-based methodology with respect to the disciplinary variants of enhancement. In this way it will be possible to clarify by negative approximation – that is what we DO NOT intend to say or deepen – the nucleus of the few applicative exemplifications/modalities of human enhancement taken into consideration. We shall also be able to trace the risks of dematerializing trans-human turning-points, if there are any, and endogenous hazards, in the use of these, somewhat limited, exemplifications.

Otherwise, there can be the risk of a latitudinary and omnivorous vagueness of the concept of human enhancement. This may be such as to cover only superficially the points of view and predicative expressions and semantic descriptors typical of many fields of knowledge and experience. Such a 'poor conceptual infinity' is more than incumbent, and such as to become a paralyzing inescapable certainty, with risks of inducing indifference and saturation in those dealing with the theme of the *overcoming* of conditions (any whatsoever) in which we humans find ourselves, starting from the most varied interests.

Some mention of the 'phantasmagorical palette' of disciplinary variations: for the applied sciences, the prefixes 'neuro', 'bio', 'psycho', 'pharma' and 'nano' among others, are those which characterize the sciences which are at the forefront in setting their sights on an enhancement of the human properties and qualities that are object of their own field of observation and intervention. This is without forgetting massive use within the context of aesthetic plastic surgery and cognitive sciences. Furthermore, for the pure sciences, the German term of *Optimierung*, which in some contexts is used as synonym of *enhancement*, in mathematical jargon means the search for the best standard for a complex system. This is simply to give a minimal idea of the semantic pluriverse and misleading superimpositions that can emerge due to the 'ingenuous' and not too overcharged use of the formulation.

We shall not make reference to previous applicative modalities of the phrase 'human enhancement', but only to those, through reference to the post-human condition as mentioned above, associated with bionics, mechatronics and the spill-over effects of both, when reclassifying the subjects involved and assessing the preconditions and consequences are of implanting artificial cybernetic grafts in living organisms. The difference considered to be

conclusive is that between mechatronic implants grafted into 'peripherical' organs of the body and implants located in the centre, among others but primarily, of the ipseity and individual personality of 'us' humans: the brain. In other words, the so-called (in jargon) cyberware is the set of electronic devices (mechatronic in most cases) that are grafted with a therapeutic and rehabilitative purpose into the human organism. It can be sub-divided into not merely wearable surgical prostheses, and 'interfaces' between systems of binary codification and sections of the cerebral mass, that is between bodyware and headware.

3. Cyberware and its facets. Distinctions of a Single Genus?

We can ask ourselves whether the distinction between electronic prostheses (not merely wearable) and cybernetic interfaces is effectively necessary in terms of category or is legitimate only at a pragmatic level. That is to say, they are two phenomena located on a continuous line. There is no difference of ontological (only structural) range between bodyware and headware, once the intervention has been carried out with equipment installed in the body by means of precise surgical operations, appropriately connected and functioning – to ensure long-lasting efficiency - with *permanent* electronic arrangements and *connected to the nervous system*. Furthermore, the biological and neurological signals in question are activated in the appropriate organic sites, and the appearance or operative modality of the device that triggers them does not seem determinant, as long as it is directly or indirectly interconnected to the brain.

Let us return to the mechatronic examples: the prostheses (cochlea implants) inserted into the internal part of the ear (cochlea) involve persons who due to a serious impairment of the auditory hair cells are strongly disabled. This equipment presupposes a 'relatively' non-invasive surgical operation. In any event, the internal components of the device are implanted in the cranium and durably connected to electrodes located near the cochlea, in the internal ear. This is in order to stimulate the vibrations of the otherwise damaged ear drum which in this way enables the auditory nerve to transmit the sound information to the brain. In my opinion it is decidedly arguable to have the presumption of considering 'peripherical' the main organ used to perceive the world in the polyphony of the vibrations of the elements, of the planets, of the tides, of the melodies of composers of all cultures.

Furthermore, the mechatronic implants (bodyware) which today replace or integrate the damaged parts of our body are built to interact repeatedly with the somatic, neurological and psychic context, not only with the bio-physical substrate in which they are inserted. To the greatest extent this is true for the limbs, first of all the hand, which can no longer be seen as in the past, and this can be affirmed on account of the dynamic and profound compenetration existing between the most advanced prototypes and the psycho-physical-biographical-ideal identity of the subject, and also due to the anthropic camouflage of the organ, which is much more successful than in the past.

On the one hand, there are interfaces also here, not only in the exclusively cerebral grafts; they operate by repeatedly connecting the nerve terminals and the silicon chips that guide and govern the mechatronic limb. They enable the subject to perceive, albeit with relative discontinuity, the motion and the limb carrying out the movement as if they were the body's 'own'. The central control room of the holistic (bio-chemical-neuro-psycho-socio) plexus when interconnected and positioned, which each individual is, is always involved.

On the other hand, the 'coverings' of the limb are imagined and designed so as to have configurations that are increasingly closer to the original organ, the human skin. There will be fewer and fewer of the disquieting 'metallurgical' forms, which still disturb western citizens/users. There is still widespread deep and irrational repulsion towards the possibility of uniting human biological purity (an individual) with artefacts (a prosthesis of steel or other metal) deriving from industrial mechanical or mechatronic manufacturing. A prosthetic device *without imitative biological covering* is seen as the most unashamedly unnatural and inorganic *res extensa*, a modality of material that is so clearly artificial because it is 'machinal'[3]. Furthermore, it is the thing that is furthest away from the incorporeal and rarefied enlightenment of human thought. This is, obviously, according to the negatively hyper-reactive sensitivity regarding the 'hybridisations' between body and machine typical of the western hemisphere.

Let us recall how only a few decades ago 'external' prostheses were perceived, as visually and structurally alien and abnormal with respect to the organicity of the human body. They were perceived, in the worst cases, as hideous, and in the best cases as ignominious, both by the person who had to 'suffer' in order to stem the effects of a disabling trauma, and also by the family

---

[3] This neologism, from Latin roots, is preferred to the term 'machinic'.

and the social context. Only the most subdued and softer versions of these 'vile' devices were, let us not forget, the hooks and wooden legs of 'always negative' personages belonging to the adolescent imaginary.

Coming to terms with some of the 'abominable' nightmares produced by our deepest theological and speculative legacy should crank us up a level compared with tales, and enable us to accept a socio-cultural therapy against 'contact phobia', together with a therapy, counterbalancing the first, which opposes the 'cyberfusionistic' syndrome of the post-human variant. According to the latter, the materialistic model of the Homme Machine becomes the technophile dystopia, socially prescribed and collectively pursued, of the perfect Human Machine.

4. Bionic Beings and Cyborgs – Meaning and Usefulness of the Distinction

From this point of view, we must not underestimate the implications, also in terms of fundamental rights, for subjects involved which can originate from a *mistaken*, unsuccessful, intervention of enhancement, even worse if ascriptively imposed from outside and not reflexively accepted, as would be the case of a service conforming to the modelistic-normative and technocratic aspect of a preventive/systemic 'enhancement'. Even more serious is the case of an intervention which was not intentionally desired and pursued individually (Koops, 2013). This warning is particularly appropriate to those situations regarding beings – not only imagined, but of the near future – which are empowered by mechatronic components that are proportionately predominant with respect to the organic components: cyborgs.

It has been said that the cyborg 'lives' in the popular imaginary but moves, in accordance with a brief but dense history, in the more concrete domains of material and immaterial production, laying bare the potential and hazards of the often conflictual relationship between 'human being' and machine (Caronia, 2008). Due to its technical-manipulative viscosity and its link to cybernetic ideative and productive processes, human enhancement *should* never assume the features of *Human Engineering*, even if there is still need to translate this '*Sollen*' into a grammar, a syntax and biopolitical pragmatics of social conflicts (Bazzicalupo, 2010, Haraway, 1997). What is more, it would be catastrophic and aberrant if who pursued enhancement as an item on the political agenda were blind to both individual, biographical and contextual differences, and also to the axiological implications that inevitably derive from

initiatives of 'hybridization'. In this regard, it is the inventors, manufacturers and therapeutic practitioners involved who must firstly assume responsibility, although not exclusively. At this point we must briefly take a look at the mechatronic typologies of these forms of hybridization.

We know from some literature (Henry, 2013, Tagliasco, 1999, Haraway, 1991) and from recent successes in rehabilitative prosthetic robotics, that cyborgs ideally represent the type of anthropoid that are neither totally organic nor totally mechanical (more precisely, mechatronic), whose numerous configurations are located however along a single line. Moreover, according to Donna Haraway (Haraway, 1991), the distinction between headware and bodyware does not hold, as regards category, like, in the opinion of the author, bionics and cyber-mechatronics are *conceptually* analogous. Nevertheless, we need the categorical distinction as a *criterium* to distinguish case to case the extremely various empirical phenomena.

At the two extremes we can find, on the one hand (that of the human bionic being), the greatest extent of dominance of living tissues, with only limited electronic and bio-mechanical (prosthetic) inserts: the case of the individual rehabilitated by prostheses, or bionic. Bionics is, in fact the science of systems in which functioning is based on that of natural systems, or those that present specific analogies and characteristics with respect to them. This discipline enables the creation of artificial organs that are perfectly interchangeable with natural organs, damaged or destroyed by traumatic events.

On the other hand (that of the cyborg), we must contemplate a minimal degree of organic components, imagining an artificial contraption endowed however with the most sublime organic component, neither external nor superficial, however characterizing the human being more than any other: the brain. Only in this second case, of minimal but distinguishing organic presence, would we really have to do with a cyborg. In reality, it is considered that the two terms, bionic and cyborg are equivalent in terms of category and that they are located along a continuous line.

Bionics, more specifically, enables the creation of artificial organs that can replace or in certain cases can be more powerful than the natural components, whenever the treated subject has been undergone pathological disabling or even total destruction of the organ. In this case the operation necessarily constitutes an enhancement due to the radical nature of the damage, requiring replacement of natural components with artificial components still more sophisticated than the original ones. In fact, performing the functions of the

organism is generally possible at the cost of using 'supererogatory' technology, which, in order to obtain an essential result, must reproduce a performance at a higher level. Or perhaps this is a way of saying that scientists and medical practitioners maximize the usefulness of the operation to up the ante and get better results.

The contacts with electrophysiology and neurophysiology are what enable an increasingly better interaction between bionics and medical and rehabilitative research, aimed at restoring motor and cognitive functions. It is not by chance that from this disciplinary branch, by means of the biunivocal relation between science and imaginary, there originated already from the 1980s some very *realistic* fantasy characters, protagonists of television series – such as *The Million Dollar Man* and *The Bionic Woman* – which became cult series and which are still available on the Internet. In the world of comic strips, still earlier than that of the cinema, a precursor was Iron Man (also a super hero), who is saved from certain death by sophisticated armour made up of a plate of sensors and electromagnetic devices that prevent shrapnel from reaching his heart. In this way, the prosthesis necessary to save life becomes at the same time an instrument of unusual strength and wide-ranging potential, such as to require equally painful and radical transformations of the personality and identitarian make-up of the protagonist.

It is therefore difficult to accept a super-humanistic connotation for such characters, which, if anything, are symbolic proof of the structural and insurmountable creatural fragility that is common to the living. The brutal alternative is between a clear leap beyond the human condition of common mortals and physical and cerebral death. Human enhancement can sometimes be a necessity that is accepted *unwillingly*, not an act of arrogance or expression of delirium of omnipotence. If anything, it is symptom of a profound ontological weakness, or of unflagging, perhaps criticable, attachment to life and the world, which has been dealt with elsewhere (Henry, 2013).

One can certainly dread also in such a rehabilitative application a serious risk, innate in bionic and post-human enhancement; this is not a frenzy of omnipotence annihilating the embodiment of the identity (as for the dematerializing trans-humanism), but rather an obsession of biographical immortality, to be pursued at all costs. In this case such a syndrome would not have phantasmal characteristics of an immaterial and reticular type, but if anything of a hyperhedonistic and solipsistic sort. Moreover, it would produce socially asymmetric and unequal effects. Only the powerful of the Earth could

afford to perpetuate the numerous and expensive operations of bionic limbs, to the point of becoming complete cyborgs, no longer *perfectable, of a limited number on account of economic asymmetry and undisputable overlords of the non-cyborgs*. The perfect and classist human machine.

Also considering on the other hand the possibility of a democratization of the social good 'bionic-cybernetic enhancement', there still exists the repulsion and fear of the combination of what is human and what is artificial, of which we have spoken as regards western societies. However, these reactions do not seem to be triggered with such immediacy in another case, already widely urbanized in the (mainly) male imaginary, and of which we shall speak in the next paragraph.

5. Under the Ambiguous Sign of Sexed, Transgender and Artificial Humanoids

It can be observed how the thesis regarding the plasticity and manipulability of human corporeity terrifies many when its outcome is dystopic figuration, but realistically obtainable in the near future, of a cyber human being or perfect human machine[4].

In a second case, not categorially different from the first, which is however abhorred due to its 'symbiotic blasphemy', is that in which we hypothesize, imagine or desire a humanoid (golemic creature, or android, or cybernetic being) which is sexed and used for erotic purposes.

We are alluding to a combination of visions and sometimes unconfessed aspirations deriving partially, but not only, from middle-eastern, in particular Hebrew, mythographic tradition. Without doubt, the fact of hypothesizing the sexual use of humanoid bodies, both male and female, can find confirmation in an ancient and documented tradition of kabbalistic reflection on the usefulness and appropriateness of golems, as Idel in particular showed us in his fundamental work on the prototypical artificial humanoid (Idel, 1990, Henry, 2013). This *traditio*, however, no matter how noble its lineage from a

---

[4] A reaction of rejection that is not comparable can occur with respect to a different example; when special techniques are used for specific artistic performances directed at going beyond the *limes* between nature and artifice these are often considered as belonging to the particular case of the *cyborg*, such as the so-called body-machine performers, which is surely closest to being human given the temporaneity of the grafts and the corporeal manipulation of the artist, although it is aimed at exasperation of vision according to which organism and machine appear to the spectator as if they were in full symbiosis. Cf. Tagliasco (1999), p. 81.

mythographical and symbolic point of view, is not as influential on the imaginary as the expressed or unconscious desires of many (principally male) humans, at a global cultural level. The latter, ennobled by a distinguished western and oriental literary and cinematographic tradition, would be happy to possess an artificial geisha, a mechanical doll, like the cold, silent, untiring and luminous artificial lover of Fellini's Casanova, or rather a *woman cyborg*, made up of all the women that they could hypothetically desire, and suitable to satisfy (at least presumptively) each and every erotic and emotive need. In a passage from *Zeno's Conscience* there is a renown example which anticipates by more than a century, with respect to 'natural' feminine examples, the mental and emotive disposition which is welcoming as regards 'feminine cyborgs[5]:

> "I was sincere as in the confession box. I did not like woman in her entirety but ... in pieces! Of all I loved the feet if well shod, greatly the neck if slender or even if strong and the bosom if light, light. And I went on counting the female anatomical parts, but the doctor stopped me: 'All these parts make the woman whole.' At that point I said an important word. 'Healthy love is that which embraces a single and whole woman, comprising her character and intelligence.' Up until then I had certainly not encountered such love and when this happened it made me suffer, but it is important for me to have seen the illness where the doctor saw health and that my diagnosis came about" (Svevo, 2010, p. 39).[6]

With regard to this, in comic strip creations and science fiction there often circulates the idea of time travellers and astronauts who request the possessors of cybernetic and mechatronic knowledge to supply them with artificial

[5] It is an invasive heterodetermination as can only be the still dominant and ruling male vision, which imposes *enhancement* not only via aesthetic surgery, but also 'a' stereotypical and sexist model of perfect wife/companion, as depicted and crticized in 'The Stepford Wives' of Bryan Forbes and the recent remake.

[6] This tradition includes, for example, in the field of male desires: the fair and graceful mechanical Olimpia from the tale of Hoffman, *Der Sandmann*, the cold and perturbing robotic copy of *Metropolis* by F. Lang, the artificial creature built by Dr. Frank-N-Further as object of sexual pleasure in *The Rocky Horror Picture Show*, and the replicant Pris, of *Blade Runner*. On the side of female desire, keeping to just a few examples, we have the android lover and cohort of Barbarella and, outcome of the dark side of the imaginary, the unsettling and handsome Necron, assembled, like Frankenstein, with biological parts from his female creator, the necrophilic virago Frieda Boher, the same referred to in the first number, of the same title, the "fabbricante di mostri" (the monster maker); this Italian erotic comic strip, *Necron*, from 1981, is by Ilaria Volpe and Roberto Raviola, with drawings by Magnus). As an axiologically positive example, even though with tragic outcome, we have Yod, the *cyborg* who receives a complete sentimental and erotic education, according to the plot of Piercy, (1991) *He, She and it*. Many thanks to Paola Bora for enabling me to appreciate this novel.

duplicates of themselves, in analogy with an Amphitryon who is this time consenting, in order to discourage their partners, not contrary to the project, from having the need and opportunity to replace them with other more forbidding human sexual *partners* during their long absences from home. We are referring to androids equipped for sex (Tagliasco, 1999, p. 278).

Of another tone with respect to the previous vision, which however expresses projections, apprehension and widespread desires (not only in the Western world) which should be taken into account, is a third vision, the following.

It is that which, in the field of gender studies, and starting from the works of Donna Haraway and part of the cyberpunk literature, seizes this cognitive challenge as a chance to contrast and contest - through original alliances between socio-anthropological subjectivities and cybernetically connected configurations – the hubris of those who endorse the sacral purity of an exclusively biological origin of the born of woman in rigidly defined sexed bodies by means of mating with natural methods. The latter, the enemies of the 'not born of woman', can be considered as the renewed disciples of the ideology of *limpieza de sangre*. This requisite in itself would be a factor of ontological and moral superiority of the 'original' humans with respect to all the beings that are organically spurious, metamorphic or hybrid that can be hypothesized or that already exist from a genetic, genealogical and sexual point of view: the 'rejects', according to the traditionalists of biological and specist purity, range from artificial humanoids (golemic beings, robots and *cyborgs*) to the constellations of *transgender* humans. In this type of practical-moral attitude, an ascriptive sort based on biological-genetic and specist ideas, which numbers many pernicious examples in the history of intolerance and racism, not only in the West, it is the purity of biological pedigree, the genealogical belonging to a category, and not the good intentions or actions of themselves and interlocutors, that counts as ultimate and distinguishing division between what is acceptable and what is deplorable in the intersubjective relations among reasonable and responsible agents.

Conclusion

In what has been said so far the intent is to give definatory credibility to an enhancement that is post-human, rehabilitating and restoring, and certainly not trans-human, that is supermanistic and directed at the annulment of the corporeity and creaturality open to relationality between finite beings. It is the

meaning according to which all creatures, even more so the most powerful, must be companions in the sharing of pain, as in sharing the aim of limiting the inauspicious impact on the living and whoever is animated, and it moves within the inner-worldly and infraspecist horizon. This opening should for the sake of coherence be even wider apropos hybrid humanoids, of what in effect *we humans* are or become whenever we oppose the hardships of our destiny with the grafting onto our bodies of artificial rehabilitative or replacement devices, becoming cyborgs, closer in this to the inorganic and metallic aspects of our non-human companions (the automatons of the present, the androids of the future), in an opening that is reflexively accepted, and not merely endured (Fadini, 1999). This mutation implies, per se, a widening of the horizon of inclusion of morally qualified subjects and it will take place, perhaps, in a not too distant future. This will happen whenever there is among our interlocutors, within the pragmatic state of coexistence and social life, whatever humanoid capable of foreseeing and accepting the consequences of their choices as regards other subjects involved, however modified or altered they are with respect to a presumed original human model. Whoever is hybrid or bionic bears *written in the body* the hazardous and ambiguous burden of being the target of discriminatory practices legitimized by their not being fully *human* or of being so *unrestrainedly*, beyond what is consented. The seed of racism and xenophobia lies in learning the noxious, but typically human, taste for humiliating he who is *other*. The label can be attached to whoever one wishes, the step is short. What counts is contributing to building, politically, from now on, the conditions contrary to these deviations. They must be the widest possible conditions, of non-hegemonic or expansive conscious contextuality of legislative and decisional systems, focussing rather on the social circumstance whereby we see ourselves as subjects that already co-inhabit multiform social identities, in changeable and hybrid bodies and identitary images, in potential or latent conditions of moral and political asymmetry, which are therefore to be preventively identified and neutralized. "It does not count 'what' you are, but 'who and how' you decide to be". This could become the legislative principle of a society of humanoids - *humans*, hybrids, artificial beings – who are free and equal.

REFERENCES

Bazzicalupo, L. (2010). *Biopolitica. Un Mappa Concettuale*. Roma: Carocci.

Bernal, J. D. (1929), *The World, the Flesh and the Devil. An Enquiry into the Future of the three Enemies of the Rational Soul.* London: Jonathan Cape.

Caronia, A. (2008). *Il Cyborg. Saggio sull'uomo artificiale.* Milano: ShaKe.

Cavallarin, M., Henry, B. (2012), *Gli ebrei in Cina e il caso di «Tien Tsin». Convivenze nella Cina moderna*, Livorno: Belforte.

Coenen, C., Gammel, S., Heil, R., Woyke, A. (2010) (eds.), *Die Debatte über „Human Enhancement". Historische, philosophische und ethische Aspekte der technologischen Verbesserung des Menschen.* Transcript Bielefeld.

Fadini, U. (1999). *Principio metamorfosi. Per un'antropologia dell'artificiale*, Milano: Mimesis.

Gibson, W. (1984). *Neuromancer.* Canada: Ace Science Fiction.

Gibson, W. (1987). *Count Zero.* London: Penguin.

Gibson W (1988), *Monna Lisa Overdrive*, Victor Gollancz Ltd, UK.

Haraway, D. (1991). A Cyborg Manifesto: Science, Technology and Socialist-Feminism in the Late Twentieth Century, in Haraway, D. (ed.), *Simians, Cyborgs, and the Women: The Reinvention of Nature.* New York: Routledge, 149-181.

Haraway, D. (1997). *Modest_Witness@Second_Millennium FemaleMan _Meets_ OncoMouse,* London: Routledge.

Henry, B. (2011a). *Vorstellungwelt, Kulturen und künstliche Identitäten.* In Cerutti, F., Rudolph, E. (ed.), *Brauchen die Europäer eine Identität?* Zürich: Orell Füssli, 249-263.

Henry, B. (2011b). "Transmigrations among Imaginaries among East and West", *International Journal of Humanity and Social Sciences* I 2, 33-37.

Henry, B. (2013). *Dal Golem ai cyborgs. Trasmigrazioni nell'immaginario.* Livorno: Belforte.

Koops, B.-J. (2013). *Concerning 'Humans' and 'Human'Rights. Human Enhancement from the Perspective of Fundamental Rights*, in Koops, B.-J., Lüthy, C. H., Nelis, A., Sieburgh, C., Jansen, J.P.M., Schmid, M. S. (eds.),

Engineering the Human. Human Enhancement Between Fiction and Fascination. Berlin-Heidelberg: Springer, 165-182.

Labanca, N. (2012). *Prefazione,* in Cavallarin, M., Henry, B. (eds), *Gli Ebrei in Cina* cit. Moravec, H. (1988), *Mind Children. The Future of Robot and Human Intelligence.* Cambridge, MA: Harvard University Press.

Piercy , M. (1991). *He, She and it.* Chicago: Middlemarsh.

Rivera, A. (2005). *La guerra dei simboli. Veli postcoloniali e retoriche dell'alterità,* Dedalo: Bari.

Svevo, I. (2010). *La coscienza di Zeno.* Roma: Newton.

Sieben, A., Sabisch-Fechtelpeter, K., Straub, J. (2012) (eds.). *Menschen machen. Die hellen und die dunklen Seite humanwissenschaftlicher Optimierungs-programme.* Bielefeld: transcript.

Tagliasco, V. (1999). *Dizionario delle creature fantastiche e artificiali.* Milano: Mondadori.

Takanishi, A. (2008). *Humanoid Robots as Tools for Scientific Study of the Human Behavior* (unpublished paper manuscript).

Woyke, A. (2010). Human Enhancement und seine Bewertung. Eine kleine Skizze. In Coenen, C., Gammel, S., Heil, R., Woyke, A. (eds.), *Die Debatte über Human Enhancement,* Bielefeld: transcript, 21-38.

Transhuman Perfection: The Eradication of Disability Through Transhuman Technologies

David-Jack Fletcher [†]
david-jack.fletcher@mq.edu.au

ABSTRACT

This paper examines transhuman technologies that seek to eradicate disability - primarily prostheses and implants. While most would agree that disability denies individuals the same quality of life as those deemed "abled," this eradication ultimately relies upon secular humanist notions of the perfect human. Transhuman technologies hold obvious implications for the human body, however they also hold implications for what it means to be an acceptable body; ultimately these technologies aim to create the perfect human by eradicating the disabled Other. This paper uses these notions to question concepts of "hierarchies of life," at which disabled individuals are most commonly moved towards the bottom, or at the very least considered nonhuman. This article seeks to provide alternative theory to the eradication of disability, which states that these individuals may not have the same mode of existence, but that their mode/s are just as valid as those lived by "abled" individuals through an examination of Braille.

Introduction

The human body as a site of inquiry is not a contemporary concept, and notions of what classifies as a human body has largely influenced biopolitical regimes and sovereign power. Biopolitical discourses that culminated in the Nazi eugenics regime during World War II held the belief that specific types of bodies were inferior to others, and ultimately classified as inhuman, which resulted in the liquidation of countless individuals under the rubric of racial hygiene. Nazi eugenics is an extreme example of both the sovereign power over

---

[†] Macquarie University, Sydney NSW, Australia.

life and death, and a quest for corporeal perfection; more subtle examples can be seen in contemporary Western society, such as the treatment of disabled individuals. Many forms of eugenics were discriminatory by their very nature, operating within a system of exclusion. Indeed the "old form of eugenics discriminated against the disabled and less intelligent by forbidding them to have children" (Fukuyama, 2002, p. 159). While this is no longer the case and those deemed disabled are quite freely allowed to procreate, disabled individuals are encouraged to conform to corporeal hegemony in various ways, which can be seen as resurgence of eugenic regimes. While the Nazi regime was primarily preoccupied with issues of racial hygiene, eugenics in contemporary society has transgressed to issues of corporeality and *genetic* hygiene; at the core of both, however, are greater issues of equality and hegemony that position specific bodies as superior to others.

Of concern in this article is the disabled "human" body – that is, those with "cognitive and physical conditions that deviate from normative ideas of mental ability and physiological function" (Mitchell & Snyder, 1997, p. 2) – for its perpetual recognition as somehow "less than" human, and the subsequent marginalisation and disenfranchisation of these individuals. The marginalisation of these individuals demonstrates an exclusionary system still in operation in contemporary society, igniting the notion that eugenic regimes have evolved and are applicable to discourse on disability. This article will scrutinise the addition of what I will posit as transhuman technologies – such as implants and prosthetics – to disabled bodies that seek to eradicate perceived physical and/or psychological deficits, and the implications these technologies hold for notions of acceptable human bodies. It is important to note here that disability is not the only issue to be dealt with regarding these sorts of technologies, as aging and the quest for immortality can be highly linked with disability studies and transhumanism; however, this paper will specifically focus on the issue of disability, as to examine anti-aging and immortality would require extensive space.

Questions of what constitutes a human body are perhaps of greater importance now than ever before, due to the addition of advanced technologies to these bodies. Quite often it is often overlooked that "from the very beginning ... the body is always already intextuated and instrumentalised by a series of technologies" (Pugliese, 2011, p. 946), and indeed can be viewed as the original prosthesis able to be manipulated (Hayles, 1999). The possibility, then, of adding technology to any human body infringes upon notions of that

body remaining human, but rather becoming transhuman – and, eventually, posthuman. Nick Bostrom identifies transhumanism as a movement that seeks to enhance the human through technological advancement, "like genetic engineering and information technology, and anticipated [technologies] such as molecular nanotechnology and artificial intelligence" (Bostrom, 2003, p. 493). This article will adopt Bostrom's understanding of transhumanism, and through this definition it is possible to view both implants and prosthetics as transhumanist technologies.

Implants and prosthetics now make use of nanotechnology to integrate seamlessly into the body of the recipient, however "applications derived from nanotechnology have the potential to further marginalise those in society who are perceived as disabled" (Sheremeta, 2004, p. 51). The visibilisation of difference – both pre- and post- nanotechnological implants/prosthetics – can act as a proponent of the "lack" of a pure human body, thereby placing them further within the category of other. This article posits implant and prosthetic technologies are transhumanist in nature for their capacity to modify the human; whether these modifications are positive or negative remains to be seen. This article suggests that it is through these transhuman technologies that the eradication of disability will most likely occur, unless one acknowledges the validity of modes of existence that lie outside Western hegemony. Through an examination of the technology of Braille and recent advances in adaptive computer software, this article suggests that transhuman technologies need not "amend" these bodies, but perhaps work with them to acknowledge the validity of non-hegemonic modes of existence, and potentially disrupt biopolitical discourses that seek to eradicate disability.

1. Disabled Bodies And The Transhuman Other

The figure of the disabled is generally viewed, in Western society, as undesirable, and a site for pity and/or disgust; these bodies act as reminders to the able-bodied that things can go "wrong". Hence, the "disabled are constant reminders to the able-bodied of the negative body – of what the able-bodied are trying to avoid, forget and ignore" (Fitzgerald, 1998, p. 152). The image, then, of the able-bodied, becomes normalised in a sense, and "normality [becomes] an assumed state which reproduces itself through a visual registry that engenders bodily integrity as self-evidently visible" (Sullivan, 2005, p. 332). Abnormality, too, however, makes itself visible in the sense that these

states are recognised as deviations from the assumed "normal" state. Recognising disability as an undesirable characteristic, and potentially even as a threat to the mythical "pure human" can ultimately position the human as fragile; it is possibly this fear of fragility that drives the human to pursue the dream of perfection through technology. One can argue that the blatant shunning of the disabled Other by homogenous society has led to the development and deployment of transhuman prosthetics; in an effort to both restore the body to a state of normalcy, and furthermore to achieve the ultimate humanist goal, which is that of the perfect human.

Moreover, the potential of "disability" threatens notions of the pure human, of what is "proper to man" (Derrida, 2002, p. 409), and opens a line of inquiry into the anthropomorphic nature of the human. Derrida's notion of anthropomorphism alone enables one to question the supposedly clear lines between human and nonhuman, thereby resisting the concept that humans are unique. When advanced technologies are added to this mix, the lines are further blurred and the human must reconsider its own existence in relation to other species, and acknowledge that the human may not be the sole possessor of qualities such as essence. Martin Heidegger, in *Being and Time* (1967), understands and positions essence as the unique human capacity for consciousness and rationality, which is a humanist paradigm outlining notions of the human as pure; the notion is itself a mode of disenfranchisation that de-values nonhuman species which ultimately seeks to preserve the human's position atop a (hu)man-made hierarchical species structure. It can be understandable then, that the shape of contemporary technology aims to eradicate the threat of the potential disabled figure via the addition of transhumanist prosthetics.

"Transhumanists are lovers of life who recognise that the limitations of the human condition may be overcome through the technology of the future" (Young, 2006, p. 41), and as such, in transhumanist discourse, the manipulation of the body through prosthetics is generally seen as desirable for its capacity to enhance the imperfect body. In relation to the disabled body, however, this addition is seen as therapeutic rather than for enhancement measures. The obvious difference here is between therapy and enhancement; at the heart of the therapy/enhancement dichotomy lie notions of the "normal" body, and it is "the idea of 'normal' [...] that sets the standard around which bodies are evaluated, regulated and are even permitted to materialise" (Karpin & Mykitiuk, 2008, p. 414). This notion automatically places disabled

individuals in opposition to hegemonic normalcy, positioning these bodies as Other based on what is perceived as either physical or psychological lack.

With the addition of transhuman technologies such as prosthetics, a series of new questions arise surrounding the normal body and, indeed, levels of humanity. The ultimate goal, then, of both therapy and enhancement, and the resulting treatment of those deemed disabled, can be seen as the reinforcement of the notion of the pure human and its unique qualities of essence, consciousness, rationality "and reason, which 'distinguishes us from the beasts', [which] also confers upon the human being the power to tell the difference between itself and its non-human others" (Badmington, 2004, p. 8). The unfortunate circumstance for disabled individuals is that each of them has been grouped into this category of nonhuman by outlining supposed similarities between the human perspective of the animal and the hegemonic perspective of the disabled; indeed, it has been stated that "human mental faculties like consciousness and creativity rely to a huge extent for their development on the stimulation received from the environment" (Pepperell, 2009, p. 131). Pepperell relies on this assumption to disenfranchise both blind and deaf individuals, concluding that these particular bodies are unable to interact wholly with their environment, and as such, are not fully conscious. Maurice Merleau-Ponty describes consciousness "as the possession of an object of thought or as transparence to itself" (Merleau-Ponty, 1974, pp. xv-xvi), which seems not to discount disabled individuals as conscious. Merleau-Ponty does, however, assert that phenomenology – early studies of consciousness – incorporates the corporeal existence as the most significant aspect of consciousness, claiming that it is through his body that he experiences the world. Perhaps in this instance, as one of the pioneers of consciousness studies, Merleau-Ponty has deployed notions of the disabled as in- or non- human based on the physical lack.

Pepperell uses this denial of consciousness to further argue that only humans have the quality of consciousness and therefore blind and deaf individuals must not be fully human. In similar fashion, the dehumanisation of disabled individuals – and the subsequent comparison to the animal – can be viewed as a contemporary reinscription of the Agambenian notion of bare life; that is, those whose lives are viewed as expendable under sovereign power, which Michel Foucault describes as "essentially a right of seizure: of things, time, bodies, and ultimately life itself" (Foucault, 1990, p. 136); and as Athena Athanasiou writes, "the subjugation of human life and death to

biopolitical sovereignty comes to be what is at stake in modern technology" (Athanasiou, 2003, p. 136). That is, the eradication of disability via transhuman technology can be seen as a model of the sovereign power over the right to let die.

It is perhaps this sovereign power that initiates a model of exclusion, as "bare life has the peculiar privilege of being that whose exclusion founds the city of men" (Agamben, 1998, p. 13). The notion of exclusion, as stated earlier, can be seen as deployed in contemporary society in a variety of discourses, though perhaps most visibly through the reduction of disabled individuals into nonhuman entities. In this way, then, "the disabled have been particularly fortunate beneficiaries of the age of intelligent machines" (Kurzweil, 2000, p. 58), particularly in relation to advanced prosthetics as "upgrading" these individuals to more than simply bare life, "for those with missing or disfigured parts, passing as able bodied is important for social as well as physical functioning" (Hogle, 2005, p. 706), which is demonstrative of biopolitical regimes of homogeneity and hegemony. With this discourse of normality, one is able to make the argument that prostheses used as a therapeutic measure in disabled individuals acts as a pathway to the restoration of the full body. Perhaps, then, the addition of prosthetics enhances the state of consciousness to these individuals.

Despite any addition of, for example, a prosthetic limb, however, it can be argued that these bodies are not reinstated as "fully" human, as "visible prosthesis ... reminds us of the way in which prosthetic culture is enveloping all of society, not only the disabled and the disfigured" (Mitchell & Snyder, 1997, p. 86); this in conjunction with the fear of losing our human essence is perhaps what renders post-prosthetic individuals as just as nonhuman as they were pre-prosthetic. Thus, the "alignment of disability with fears of the inhuman" (Wills, 1995, p. 28) does not only extend towards animals, but also towards technological nonhuman species, such as the image of the cyborg. The overtly robotic nature of visible prostheses – particularly limbs – promotes the cyborgic image, thereby further threatening the mythical pure human that secular humanists so desperately try to preserve. However, as Mervyn Bendle notes, "humanity is merely a temporary stage along the evolutionary pathway" (Bendle, 2002, p. 48), suggesting that the addition of robotic prostheses, and indeed, various other forms of technology, is nothing to be afraid of and is, in fact, part of the natural evolution of the human. Despite theorists like Bendle providing a techno-progressive view, and theories by N. Katherine Hayles

(1999), Donna Haraway (1985), and Jill Didur (2003) that suggest the human has always existed in relation to technology and therefore has always been cyborgic in nature, the overall attitude of this sort of technological intervention on the human body is still at odds with transhumanist perspectives that argue for the adoption of these technologies for the purposes of both enhancement and therapy.

The hegemonic view that these technologies render the recipient as less than human places individuals – particularly those deemed disabled – in a Catch-22 dichotomy, or a no-win situation; even with the addition of prosthetic limbs which aim to restore the body to a state of hegemonic normalcy, these bodies still remind the "abled" of what can go wrong, and are a firm embodiment of an "impure" body. Regardless of the reinforcement of biopolitical hierarchies present within this dichotomy, these bodies are further positioned as Other and, abysmally, as nonhuman in the sense that they are somehow recognised as more machine than human; rendering these individuals as the transhuman Other.

2. Beneath the surface: Implants and genetic hygiene

According to Ad Bergsma "we may gain the power to redesign the human body and mind" (Bergsma, 2000, p. 403), and certainly with the development of transhuman prosthetics, this has proven to be accurate. However, it can be seen that prosthetic technologies indeed act as a proponent for the eradication of disability, while simultaneously further producing the recipients of these technologies as the transhuman Other – either way, these individuals are not seen as fully human, reinforcing biopolitical hierarchies of life. Comfortably atop this hierarchy lies the mythical pure human, though its position becomes increasingly unsteady with the deployment of advanced implants and the development of genetic hygiene techniques, which have aided the re-deployment of notions of the human. In relation to these same hierarchies, various modes of existence that deviate from hegemonic norms are thus invalidated. These implant technologies, too, may be seen as resurgence of eugenic regimes, and certainly with the development of advanced nano-implants and the potential for designer babies, it doesn't take a stretch of the imagination to envision the deployment of these technologies as a reinscription of biopolitical hierarchies and genetic hygiene, and ultimately, the eradication of disability.

The eradication of disability can be seen as a reinvigoration of eugenic regimes – or to use a term by James Hughes, "neo-eugenic" (Hughes, 2009, p. 16) – that aim to move the human towards a state of perfection; as discussed in the introduction to this article, an extreme example of this form of eugenics was the culmination of the Nazi death camps, the focus of which was the racial hygiene of Germany and other parts of Europe. The resurgence of eugenics in contemporary society is primarily focused on genetic hygiene; that is, the perfection of the human through the manipulation of genetic material, as demonstrated through the development of designer babies. As well as the intervention of genetic material, the use of nanotechnology to develop implants that alter and enhance the human can also be seen as yet another form of eugenics. While the extremity of Nazi eugenics faded some time ago, the eradication of disability can be viewed in a similar fashion when we articulate mass annihilation "in terms of mechanical economy in the age of technological reproduction; the concentration camp is cast, at a stoke, as an assembly line of decorporealisation, a technological project whereby the natural world is reduced to a 'standing-reserve' or raw material" (Athanasiou, 2003, p. 134). Just as the Jews were annihilated in the hopes of creating a "master race", so too are disabled individuals, though rather than rounding disabled people into concentration camps, technologies, such as those deemed nano – specifically implants and genetic manipulation – are dispersed upon these bodies to alter, improve, enhance, and ultimately to erase undesirable deficits most commonly seen as disability.

While implants and genetic hygiene are two very different techniques used to eradicate disability, these technologies do share an important thread, which is what I will call on for this paper; both implants and genetic hygiene aim to promote and enhance desirable traits and/or characteristics in human subjects, thereby perpetuating hegemonic norms and acceptable bodies. Genetic hygiene does this through technologies such as genetic manipulation known to Simon Young as "superbiology, [which] will enable us to enhance our minds and bodies beyond the limitations of the human condition" (Young, 2006, p. 21). The notion of designer babies ensures that undesirable characteristics are literally bred out of future generations through genetic manipulation; implants, however, arrive after a body has been created and deemed disabled, undesirable, nonhuman, or some other form of Other. The most notable implants are, arguably, neural implants derived from nanotechnology, for the fact that these implants may directly impact notions of consciousness as

discussed earlier in this paper. These technologies are not without merit, of course, and in some ways should be celebrated for their capacity to shift a recipient's quality of life to a more standardised notion of what is inferred by "quality". For example, "deep brain stimulator implants, are a remarkable therapy that relieve the tremor, rigidity and bradykinesia of Parkinson's disease by manipulating basal ganglia activity" (Donoghue, 2002, p. 1085). While there is no denying that this technology will allow people with Parkinson's to enjoy a different lifestyle, and the validity of these technologies are not being disputed per se, this paper is elucidating the notion that these technologies are necessary; that different modes of existence must be eradicated. Through the use of these advanced technologies, contemporary Western society is perpetuating hegemonic discourse not only on acceptable bodies and biopolitical hierarchies of life, but also on what constitutes the human.

The integration of nanotechnology onto, and furthermore into, the body, is potentially what drives questions of the human – for some, it becomes uncertain just how human an individual is if they have a network of wires throughout their body, a neural chip, or a robotic limb. Despite humanist technophobia, research into such technologies continues, and "nanotechnology has been prophesised to accomplish almost anything called for by human desires" (Milburn, 2002, p. 262). Indeed, the corporeal enhancement capabilities of these technologies seem almost endless. It seems to make sense, then, to derive from these technologies applications that seek to enhance those deemed disabled. However, as briefly discussed above, when these technologies are deployed onto non-hegemonic bodies, the application is then seen as therapeutic and/or restorative in nature, rather than enhancing. This may be due to the very nature of the therapy/enhancement dichotomy; the distinction between therapy and enhancement is "commonly made between interventions that are therapeutic in their intent, used to treat disease or disability, and interventions to enhance or improve on normal species function or to bestow entirely new capacities, non-health related improvements" (McGee & Maguire, 2007, p. 293). This definition reinforces Western hegemonic discourse, which becomes problematic when we acknowledge that applying transhuman technologies to a disabled body for therapeutic means positions these individuals further as Other, abnormal, and ultimately, nonhuman.

Positioning disabled individuals in this way acts as "a reminder that the body proves no less mutable or unpredictable than the chaos of nature itself"

(Mitchell & Snyder, 2001, p. 126); hence the importance to hegemony of applications derived from nanotechnologies – specifically implants and genetic manipulation – becomes increasingly clear. Applications from implants and genetic manipulation perpetuate the ideal that certain bodies are more acceptable than others, not only invalidating groups of people – for example, those deemed disabled – but also invalidating their entire mode/s of existence.

3. The Transhuman Other as Valid

Thus far, this paper has addressed the fact that transhuman technologies aim to eradicate disability and enforce Western hegemonic discourse on acceptable bodies. What has been of focus is the intent to eradicate non-hegemonic modes of existence, implying that these technologies must *overcome* certain deficits rather than *accommodating* perceived deficits. Rather than acknowledging the validity of non-hegemonic bodies and their subsequent modes of existence, transhuman technologies seek to create a master race, just as eugenic regimes have done throughout history. This is not to say that there doesn't exist certain technologies that aim to work with a person's perceived disability, as some forms of technology have been adapted and developed to do precisely that; one prime example is that of Braille, another is computer software designed for those with vision impairments. Despite the lack of an interface and advanced computer software and hardware, Braille is nonetheless a transhuman technology. The technology of Braille is written into the lives of blind individuals and rather than attempting to remove the disability through the use of advanced technology, Braille accompanies blind individuals throughout their daily lives simultaneously allowing them to continue to be blind without forcing hegemonic discourses onto, and into, their bodies and minds.

   In this sense, then, Braille acts as a form of enhancement rather than therapy, opting to exist on the counter-side of the therapy/enhancement debate than the usual placement of these sorts of technologies. As in common knowledge, people with perceived deficits "have not only been constructed as 'Other', but frequently as 'the Other' of 'the Other'. People with disability are marginalized even by those who are themselves marginalized" (Clapton & Fitzgerald, 1997, p. 1), so while scientists develop bionic eyes and restore sight, it is important that the technology of Braille remain for those who do not wish to partake in genetic engineering or nano-upgrades. Especially because,

for some, their "disability has become an essential part of their identity and genetic engineering thus challenges the worth of their own sense of self" (Fitzgerald, 1998, p. 160), so any attempt to alter their mode of existence is likely to be met with fear. The notion of altering modes of existence through either therapy or enhancement lends itself to the premise of eugenics – as discussed throughout this paper – but also to that of mastery, which is "the extent to which one regards one's life chances as being under one's own control in contrast to being fatalistically ruled" (Pearlin & Schooler, 1978, p. 5). Eugenics and mastery almost go hand in hand, however the notion of mastery if perhaps the driving force behind eugenic research; the Nazi's focused on the mastery of race, and now it seems that contemporary society is focused on the mastery of disability, thereby challenging any assumption that disabled individuals have any chance of being their own "masters", and indeed, governing both their own bodies and minds.

Braille is precisely the type of technology that enables this sense of mastery within members of the blind population, as do adaptive technologies for the blind and the visually impaired such as JAWS, PEARL, SARA, the MAGic large print keyboard, and portable Braille displays. The fact that these technologies even exist implies that the quest for perfection is not unanimous and that non-hegemonic existences are viewed as valid by at least some, however the duality of these technologies must not escape this field of inquiry. These technologies at once operate as therapeutic in the sense that they aim to allow the recipient/s to achieve a sense of hegemony, while simultaneously operating as an enhancement technology in the sense that these technologies are, in a way, linked to these bodies, altering their corporeal experience with their surroundings. This is not to suggest that these individuals require enhancement with their surroundings, however according to the description by McGee and Maguire (2007) earlier in this paper, enhancement suggests improving function and non-health improvements. Computer software like JAWS does not aim to improve the health of blind individuals, but to allow them to engage with documents and literature by reading aloud to the blind individual what is on their screen. Despite the dual functionality of these technologies as operating as both therapeutic/enhancement, they still perpetuate the mastery of the body as the recipients of these technologies have first and foremost *chosen* to utilise them, and furthermore, have chosen the extent to which they engage with these technologies. What is the most important aspect of these technologies, however, is that they acknowledge

blindness as an acceptable mode of existence and works with, rather than against, the disability.

Even with this acknowledgement, many of the blind community would rather possess the sense of sight, and "by appealing to the vast majority of disabled who strongly support enabling cures and prosthetics, progressives can marginalize the few but vocal radical disability activists who reject enhancing technologies as neo-eugenic" (Hughes, 2009, p. 16). The reinscription of eugenic regimes is highly problematic for the perception of non-hegemonic bodies as valid and seems like nothing more than the mass disenfranchisation of clusters of non-hegemonic bodies; and ultimately, the invalidation of different modes of existence for the goal of perfecting the human race. Furthermore, transhuman technologies reinforce notions of biopolitical hierarchies of life through this process of disenfranchisation. It becomes supremely important, then, that modes of existence that lie outside hegemony are acknowledged as valid, which is what the technology of Braille and adaptive computer software seek to achieve; unfortunately, there may not be enough of these technologies to allow for the acknowledgement that I propose, with the dominant goal in Western society being that of perfecting the human.

Conclusion

Notions of disability are largely based on presupposed ideological frameworks of what constitutes the "human" – furthermore, the "whole human" – particularly regarding perceived understandings of normalcy. This paper has drawn on historical biopolitical hierarchies of the human – specifically, that of eugenic regimes – and the perpetuation of these hierarchies in contemporary Western society through the disenfranchisation of disabled individuals; ultimately this treatment of the disabled amalgamates into the positioning of them as somehow nonhuman, as "less than", based on hegemonic notions of lack.

The figure of the nonhuman thus perpetuates humanist paradigms of the perfect human, the quest for immortality, and a deep-seated desperation to assume acceptable bodies. With the progression of advanced technologies, which this paper has posited as transhumanist in nature, it can be argued that this dream of perfection has led to both the development and deployment of these technologies; indeed, the use of these technologies are nothing more

than extensions of humanist paradigms that sought to eradicate modes of existence that deviated from presupposed hegemonic norms.

The eradication of disabled individuals through transhumanist technologies, such as implants and prosthetics, operates within a dichotomous structure of ideological values. They at once enhance "normal" human bodies and provide therapy to those deemed Other, which perpetuates notions of acceptable bodies and biopolitical hierarchies. The aim of this paper has been to demonstrate that the Other, and ultimately the nonhuman should not be discounted, disenfranchised and invalidated, but have their modes of existence acknowledged as just as valid as those deemed hegemonic. This paper, then, has argued for the acknowledgement of disabled individuals as valid and worthy, and advocated for the removal of biopolitical hierarchies of life that dictate and govern how bodies are viewed, and how they are viewed as disabled.

Certain technologies have been developed to adapt to specific disabilities, rather than overcoming them or removing them entirely. Disabilities such as vision impairment/blindness utilise the technologies of Braille and adaptive computer software to accommodate the deficit of the user. These technologies, then, impose enhancement qualities upon the user, rather than therapeutic qualities, which challenge the therapy/enhancement dichotomy and works to validate these bodies, despite their deviation from hegemony. The existence of these technologies is quite significant for the (re)-construction of hegemonic norms, as the recalibration of what it means to own an acceptable body may very well lead to a re-examination of the human itself; specifically in relation to other nonhuman species such as animals and cyborgs. The fact that technologies exist to accommodate certain disabilities and enhance the individuals rather than provide therapy to them initiates the notion that disabled individuals are indeed just as conscious as those deemed abled, which further disintegrates abled/disabled binaries as well as – and more importantly –human/nonhuman binaries. Technologies that operate in this way, then, have the potential to eradicate the need to create the perfect human, and encourage the notion that non-hegemonic bodies need not assimilate in order to have their mode of existence seen as valid.

REFERENCES

Agamben, G. (1998). *Homo Sacer: Sovereign Power and Bare Life*, Stanford: Stanford University Press.

Athanasiou, A. (2003). Technologies of Humanness, Aporias of Biopolitics, and the Cut Body of Humanity. *Differences: A Journal of Feminist Cultural Studies 14*(1), 125-162.

Badmington, N. (2004). *Alien Chic: Posthumanism and the Other Within*, New York: Routledge.

Bendle, M.F. (2002). Teleportation, Cyborgs and the Posthuman Ideology. *Social Semiotics 12*(1), 45-62.

Bergsma, A. (2000). Transhumanism and the wisdom of old genes is neurotechnology as source of future happiness?. *Journal of Happiness Studies 1*(3), 401-417.

Bostrom, N. (2003). Human Genetic Enhancements: A Transhumanist Perspective. *The Journal of Value Inquiry 37*(4), 493-506.

Clapton, J., and Fitzgerald, J. (1997). The History of Disability: A History of Otherness. *New Renaissance Magazine 7*(1), 1-3.

Derrida, J., and Wills, D. (2002). The Animal that Therefore I Am (More to Follow). *Critical Inquiry 28*(2), 369-418.

Didur, J. (2003). Re-embodying Technoscientific Fantasies: Posthumanism, Genetically Modified Foods, and the Colonization of Life. *Cultural Critique 53*(Winter 2003), 98-115.

Donoghue, J.P. (2002). Connecting Cortex to Machines: Recent Advances in Brain Interfaces. *Nature Neuroscience 5*, 1085-1088.

Fitzgerald, J. (1998). Geneticizing Disability: The Human Genome Project and the Commodification of Self. *Issues in Law & Medicine 14*(2), 147-163.

Foucault, M. (1990). *The History of Sexuality: An Introduction*, New York: Knopf.

Fukuyama, F. (2003). *Our Posthuman Future: Consequences of the Biotechnology Revolution*, New York: Picador.

Haraway, D.J. (1985). *A Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s*, San Fransisco: Center for Social Research and Education.

Hayles, N.K. (1999). *How We Became Posthuman: Virtual Bodies In Cybernetics, Literature, and Informatics*, Chicago: University of Chicago Press.

Heidegger, M. (1967). *Being and Time*, London: Blackwell.

Hogle, L.F. (2005). Enhancement Technologies and the Body. *Annual Review of Anthropology 34*, 695-716.

Hughes, J. (2009). *Progress in Bioethics*, MIT Press: Massachusetts.

Karpin, I. and Mykitiuk. R. (2008). Going Out On a Limb: Prosthetics, Normalcy and Disputing the Therapy/Enhancement Distinction. *Medical Law Review 16*(3), 413-436.

McGee, E. M., and Maguire Jr., G.Q. (2007). Becoming Borg to Become Immortal: Regulating Brain Implant Technologies. *Cambridge Quarterly of Healthcare Ethics 16*(3), 291-302.

Merleau-Ponty, M. (1974). *Phenomenology, language and sociology: selected essays of Maurice Merleau-Ponty. Translated by John O'Neill*, Ontario: Heinemann Educational.

Milburn, C. (2002). Nanotechnology in the Age of Posthuman Engineering - Science Fiction as Science. *Configurations 10*(2), 261-295.

Mitchell, D.T. and S.L. Snyder. (1997). *The Body and Physical Difference: Discourses of Disability*, Michigan: University of Michigan Press.

Mitchell, D.T. and S.L. Snyder. (2001). *Narrative Prosthesis: Disability and the Dependencies of Discourse*, Michigan: University of Michigan Press.

Pearlin, L.I. and Schooler, C. (1978). The Structure of Coping. *Journal of Health and Social Behavior 19*(1), 2-21.

Pepperell, R. (2009). *The Posthuman Condition: Consciousness Beyond the Brain*, Bristol: Intellect Limited.

Pugliese, J. (2011). Prosthetics of law and the anomic violence of Drones. *Griffith Law Review 20*, 931-961.

Sheremeta, L. (2004). Nanotechnology and the Ethical Conduct of Research Involving Human subjects. *Health Law Review 12*(3), 47-56.

Sullivan, N. (2005). Integrity, Mayhem, and the Question of Self-demand Amputation. *Continuum: Journal of Media & Cultural Studies 19*(3), 325-333.

Wills, D. (1995). *Prosthesis*, Stanford: Stanford University Press.

Young, Simon. 2006. *Designer evolution: a transhumanist manifesto*. New York: Prometheus Books.

Anthropological Arguments in the Ethical Debate about Human Enhancement[*]

Jan-Christoph Heilinger[†]
heilinger@lmu.de

ABSTRACT

The paper discusses the role of anthropological arguments in contemporary ethics as exemplified in the current debate about biotechnological human enhancement interventions. Anthropological arguments refer to a normative conception of what it means to be a human being and are highly contested in contemporary moral philosophy. Most often they are promoted to constrain the ethically acceptable use of enhancement technologies. I argue that anthropological arguments can play a fundamental and important role in assessing the moral qualities of enhancement interventions, but only if their normative justification and their specific content are properly determined. I offer an account how to do so, based on the contractualist and pragmatist ideal that all those who are affected by a decision of normative relevance should be included in what I call a "quasi-democratic deliberative process". However, given that they stand in need of wide agreement, anthropological arguments resulting from such a process will be rather minimal in content. In the exemplary debate about human enhancements they hence turn out to be widely – though not fully – permissive and unable to justify a restrictive stance towards enhancement interventions.

Introduction

This paper discusses the role of anthropological arguments in contemporary ethics, as they can be found for example in the paradigmatic debate about

biotechnological human enhancement interventions. Human enhancement interventions are understood as biotechnological interventions in the human organism that aim at altering human physical or mental functioning in healthy individuals, use sophisticated technology and intervene with a certain depth in the human organism.[1] This relatively broad definition allows first to see the bigger picture and allows for a general evaluation of such enhancement interventions. In a second step, it becomes necessary to focus on concrete interventions in individual cases – such as e.g. genetic engineering, mood-enhancement by psycho-pharmaceutical means or brain-machine interfaces to improve human capacities of interacting with a computer system. Only in a broader perspective, however, it is possible to identify common aspects of anthropological relevance that are connected with the problem of altering human beings with biotechnological enhancement interventions.

Anthropological arguments are understood as a class of normative arguments that rely on a normatively charged understanding of what it means to be a human being or of "human nature". Such arguments are highly disputed in the current debate, to the degree that some claim they should be eliminated from the debate altogether (Buchanan, 2009). I distinguish questionable from more convincing forms of anthropological arguments and conclude that – while in their weaker form, anthropological arguments are rightly rejected – the stronger anthropological arguments are able to play an "elementary" roll in the moral debate about enhancements. They are elementary in two ways: First, in the sense of being fundamentally important, insofar as they provide a basic orientation for human self-understanding and sketch in broad lines what is seen as human. And second, insofar as they remain often implicit and function most of the times only in the background of the routine ethical discussions about concrete ethical problems. In the latter case, anthropological arguments can be found only "trace elements" in a comprehensive moral assessment.

Ultimately, an explicit debate is needed about the essential question, which aspects of "being human" are to have normative relevance. In the moral debate

---

[1] Obviously, this definition is somewhat vague. However, a clear and stable definition seems dispensable since the distinction that is of interest here – between morally acceptable and morally questionable interventions into human beings – does not presuppose an exact definition of enhancement. For a detailed discussion of the possible definitions of enhancement interventions, cf. Heilinger (2010, pp. 59–101).

about human enhancement interventions, a general rejection of anthropological arguments hence is not an option.

1. Different Approaches To Ethical Judgements About Human Enhancement

In the field of applied ethics one finds by now a rich set of established tools for assessing the moral quality of specific problematic interventions. The basis of these established standards of evaluation are manifold: they lie in the multifaceted history of moral theory with such different traditions as consequentialist and deontological approaches, theories of justice, virtue theory and others. In the context of applied ethics there is an increasing tendency not to focus on one moral approach alone. Instead different aspects, visible from the perspective of different theories, are combined in a pluralist way to allow for a comprehensive assessment of a given problem (e.g. Beauchamp &Childress, 2009).

In the debate about human enhancement interventions the situation is similar. Several competing and complementary approaches contribute to the moral evaluation of enhancements. Among them figure most prominently (a) risk assessments, (b) concerns for justice, (c) considerations of autonomy and pressure, and (d) anthropological arguments.

(a) Any intervention in a system as complex as the living human organism cannot be completely calculated and its consequences cannot be completely foretold. Hence also any human enhancement intervention – as e.g. interventions to improve cognitive capacities of the brain or to stop the process of ageing in human cells – carries the risk of producing something other than the intended outcomes. Increased intelligence might turn into a burden; interventions into the human genome may lead to infertility. Furthermore, beyond the individual organism, the changes brought about by human enhancement interventions may include risks and side-effects on the societal level. Since the outcomes of enhancement interventions might be not as positive as intended and since their probability of success cannot be determined beforehand, the assessment of potential risks figures prominently among the different types of ethical considerations about enhancements.

(b) Another important aspect in evaluating interventions stems from a justice perspective. After all, enhancement interventions seem to be amenities. Many people on earth do not even have access to clean water, enough food, or

medical care. Hence, every biotechnological intervention to improve human physical or mental functioning above the already high level of the healthy ones in the affluent countries, contrasts in a striking way with the need to first provide elementary goods to those worse off. But even in a narrower frame: Who in the affluent countries would have access to enhancement interventions? As such interventions would most probably not be covered by universal health care, enhancement interventions would only be open to those already better off (and this in the countries that are already better off). This unequal access might cause the social gap within a given society to widen even more. The rich and healthy would become even more healthy, perform even better in qualified and well-paid jobs, while those without access to enhancement interventions were excluded from these advantages. Also, competition for jobs between enhanced and un-enhanced individuals would appear as unfair; a society in which people live much longer would stand in need for a different pension system.

Yet, the impact of enhancement interventions for justice in society could also be different. Some have claimed that human enhancement interventions could be used to level down existing inequalities between individuals in a given society in order to increase equality and justice. The disadvantaged with relatively minor cognitive capacities could, for example, get access to cognitive enhancers to boost their performance and make them better competitors on the job market. With purposeful enhancement interventions, at least some of the existing inequalities caused by the "natural lottery" could be levelled down (Buchanan et al., 2000).

(c) Yet another standard field of assessing the ethical dimension of new technologies consists of looking at the autonomous decisions for or against some human enhancement interventions. Within the liberal framework of Western societies, informed consent of the treated person is seen as a necessary condition for any medical intervention into an individual's organism (to be dispensed only under certain, strict conditions). But still, even if one would at first glance say, that an individual made an autonomous decision to undergo a certain human enhancement intervention, there might be hidden influences exerting an indirect or subtle pressure on the putative autonomous decision. The free choice of academics e.g. to use cognitive enhancers might be based less on their free decision and more on peer-pressure exerted by a highly competitive environment in which performance and productivity are the main measures to evaluate academic achievement. Because of these concealed

influences a scrupulous assessment of the alleged autonomy of any decision for the use of enhancement interventions is urgently needed.[2]

(d) As a last main type of ethical arguments in the debate about human enhancement interventions, consider what can be called "anthropological arguments", i.e. judgements about a possible human enhancement intervention that are based on a normative understanding of what it means to be human. Assumptions about what it means to be human function as a "regulative idea", they assess actions or options for actions by comparing the (intended or real) outcome with a normative and ideal understanding of what it means to be a human being. In doing so, anthropological arguments have a prescriptive force insofar as they identify certain actions as forbidden, morally acceptable or morally good.

Anthropological arguments can appear in very different forms and are often implicit. Frequent indicators for ultimately anthropological arguments are references to "human nature" or to "human dignity", to specific ideas of a "good", a "normal" or a "typical" human life, or to the existence of "inherently human traits". Talk of human nature is the most widespread form of anthropological arguments and will be at the centre of my enquiry.

Most often anthropological arguments are used in a sceptical way to criticise human enhancement interventions. It is then said that human beings should refrain from applying these technologies, because in using them certain essentially human traits would be endangered (Kass, 1997; Habermas, 2003; President's Council, 2003; Sandel, 2007). On the other hand however, some claim that in using the new biotechnological means humans would in a pointed way execute their particularly human capacities (Bostrom, 2003, 2008). A discussion of anthropological arguments must therefore not be restricted to their prohibitive use alone but has to focus also on their permissive side.[3]

Yet, anthropological arguments from "human nature" are highly disputed in the current debate. Buchanan argues for example, «that appeals to human nature tend to obscure, rather than illuminate the debate over the ethics of enhancement and can be eliminated in favor of more cogent considerations.»

---

[2] For an analysis of the complex relation of freedom and enhancement, cf. Heilinger/Crone (in press).

[3] As anthropological arguments and their underlying assumptions about some human ideals or human perfection – in their explicit and implicit forms – are used both to support and to restrict human enhancement interventions, some confusion in the debate follows as to how to assess the strength of such arguments. An analysis of the widespread though often implicit reference to assumptions about human perfection has been provided by Roduit/Baumann/Heilinger, 2013.

(Buchanan, 2009, p. 142).[4] In the following I will defend the view, that in spite of their bad reputation anthropological arguments play de facto an important role in the ethical debates about human enhancement. Further discussion of this type of arguments is necessary, because human beings rightly do assign relevance to their self-understanding as human beings. In other words: Because it matters for human beings to conceive of themselves as human beings, a thorough discussion of the descriptive and normative components of being a human being is indispensable. The relevance of the human self-understanding is particularly salient under the current conditions, in which new biotechnological interventions may factually change what is considered to be a human being.

What is ultimately needed is a theoretical framework within which the substantial debate about the content of anthropological arguments can be lead in a well-ordered way.

2. Anthropological Arguments

While reference to anthropological arguments is common – be it implicit or explicit –, anthropological arguments vary greatly in regards to their explanation and their justification. Nevertheless, they standardly comprise of the following elements:

One should not do action A, because as a consequence of A the human trait T would be altered, and T is valuable for an anthropological reason R.

Or in a positive form:

One should do action A, because A allows preserving human trait T from alteration, and T is valuable for an anthropological reason R.

Actions A are possible human enhancement interventions (as e.g. pharmaceutical interventions in the brain-functioning; or the integration of technical devices in the human organism); T are specific human traits which are

---

[4] Others have argued that particularly the restrictive use of anthropological arguments is flawed, because "there is no plausible account of human nature that will meet the conditions necessary to support" the position of "bioconservative" authors like Fukuyama, Annas and the President's Council that see genetic enhancement interventions as a threat for human nature (McConnell, 2010, p. 415).

considered to be normatively valuable (e.g. to live on average no longer than 80 years; to have a certain capacity for remembering numbers and events; to be capable of autonomous decision making; or to have to practice hard for the acquisition of certain capabilities such as playing the piano); anthropological reasons R are different justifications to value human traits T (e.g. because it is in the God-willed set of human traits; a fixed natural endowment of humans; the naturally evolved biological optimum; or a considered consensus between those who understand themselves as human beings).

In the following I will focus on the justificatory reasons R to declare some traits T as valuable (sections 2.1. and 2.2.). Here I will distinguish between what I call "weaker", that is less convincing, and "stronger", that is more convincing, anthropological arguments.[5] After that I will suggest plausible candidates for T within the framework that provides stronger justification (section 2.3.).

2.1. Weaker Anthropological Arguments

Direct insights into the normative relevance of natural facts or properties are strongly criticised. I will present three forms of standard criticism of attributing normative relevance to natural human traits per se (cf. Birnbacher, 2006). In doing so I do not want to suggest that there are no human traits that might have normative value. I do only criticise certain ways of justifying the potential value of these traits.

2.1.1. Meta-Ethical Criticism of Anthropological Arguments

Meta-ethics deals with the ontological status of moral facts and properties, particularly in scrutinising the language in which we talk about them. Hence talking of a normatively charged "human nature" raises three meta-ethical questions about the term human nature.

A first calls attention to the ambiguity of the term human nature itself which can be used differently in quite different contexts. It is utterly unclear, what

---

[5] By "weaker" I mean that the claims of these arguments are not sufficiently explained and justified. From another perspective one could say that the weaker arguments raise overly ambitious substantive claims. Hence the arguments I call "stronger" are characterised by a better explanation and justification and achieve this also by abstaining from overambitious goals that actually weaken the weaker anthropological arguments.

exactly is part of human nature. Looking at how the term is used in practice, one can find examples for nearly all kind of human traits being included in or excluded from it. For example, some say homosexuality is not part of human nature; others say it is. Some say human nature consists of studying hard to achieve some kind of position; others say laziness is an integral part of human nature. For some it is unnatural to commit suicide; for others it is a natural thing to end their lives if they only have to expect suffering without any chance of becoming healthy again. Some conceive of murder as something against human nature; for others it is clear that aggression which may lead to murder only is a natural human trait.

These random examples are meant to show how flexible the use of the term human nature can be and how often it can merely be a projection of particular value judgements. The notion of human nature as giving support to value judgements is in these cases empty and does not fulfil its intended function to justify the normative component in any given human trait.[6]

A second form of meta-ethical criticism appears in form of the reproach to commit a "naturalistic fallacy" if attributing to any factual claim direct normative relevance. It is assumed here, that there is a gap between the realm of factual and normative propositions and no direct, logical way is leading from one side to the other. In claiming that something is the case (as in saying, something is part of human nature), one has not said anything of normative relevance yet. As Moore famously put it (Moore, 1903, § 12), any "naturalistic" explanation of the value term "good" – e.g. saying that "good" means promoting the well-being of an organism – leaves us with an "open question": Why is this (e.g. promoting the well-being of an organism") good? The same holds true for any explanation of normativity through reference to human nature: Why is human nature good?

A third form of meta-ethical criticism focuses on arguments that presuppose some kind of privileged access to this normative knowledge, e.g. if the evaluations result from God's revealed will and are perceived through inspired insight. Such "esoteric" justifications are meta-ethically not convincing, as they refer to a basis of knowledge that can not be shared by all, because it presupposes adherence to a certain religion or access to revealed

---

[6] There are more conceptual difficulties resulting from whether to conceive of human nature as of traits that have to be fulfilled in every human individual to make it fully human or whether human nature qualifies only such traits that are generally/typically/normally possessed by human beings. I will not go into these details here.

truths. Hence these insights cannot be subject to close investigation and cannot claim to be valid for non-believers.

2.1.2. Ethical Criticism of Anthropological Arguments

The ethical criticism of weaker anthropological arguments starts also from the examples given above to illustrate the ambiguity of the term human nature. Natural facts, facts about human nature or normal human traits often stand in radical contradiction to what moral theories declare to be good. It cannot be denied that human beings in all times have committed murders, exploited others, raped and cheated etc. Human beings are subject to diseases and often die very young.[7] There is no reason to declare these facts about human nature morally good simply because they are natural.[8]

2.1.3. Pragmatic Criticism of Anthropological Arguments

A last class of criticism of normative anthropological argument stems from their abuse in the past. Many of the obvious injustices among human beings have been justified with regard to "natural difference" or in pointing to a "natural position" of inferiority of some group of human beings towards others. Examples would include the oppression of women, of religious minorities, of strangers, the enslavement of others etc. All of these have been committed with some implicit evaluative assumptions about a human nature and natural order.[9]

Summing up the critique of what I call "weaker anthropological arguments": some of the claims about normatively relevant insights about human nature or what it really means to be a human being can be criticised on the basis of meta-ethical, ethical, and pragmatic reasons. Any moral claim about what human beings should be like (which human traits are particularly valuable) needs a justification that can withstand the mentioned criticisms. In

---

[7] Furthermore it is part of nature to regularly bring about droughts and famines, earthquakes and volcano eruptions, causing the death of many living beings, supporting doubt that "the natural" is in any way morally ideal.

[8] This is a critique prominently put forward by Mill in his essay Nature (Mill, 1874). – Of course there have been several attempts to declare these facts as good, most often within the tradition of theodicy.

[9] Obviously a pragmatic argument is a relatively minor argument, as it could have been for contingent reasons that these developments took place, but nevertheless I mention it here, as it sensitises for the possibly high impact of arguments from human nature.

the next chapter I will suggest a way of providing such justification for what I call "stronger anthropological arguments".

2.2. Stronger Anthropological Arguments

The above mentioned criticism showed that no direct normative claims about value and about the moral quality of particular actions result from appeals to human nature. Even if objective, universal, eternal, and scientifically proven normative claims about human nature seem to be an attractive goal for enquiry, this pursuit is a dead end. But as it still matters for human beings to conceive of themselves as human beings – as it is particularly obvious in the ethical debate about human enhancement with its frequent invocations of human nature – one cannot completely avoid thinking about a normative understanding of what it means to be a human being. The notion being human is, because it is referring to the speaker him or herself, always of a special status:[10] It is never a distant description, but includes some normative relevance. Hence it is necessary to provide a better justification for any normative claims connected with the idea of being human. As direct ways to justify the normativity of the term human being are impossible, the only remaining option consists in taking an indirect way. In the remaining part of this paper I want to sketch and illustrate this indirect way.

Some remarks beforehand: The result of what I call stronger anthropological arguments are not universal and eternal moral facts or truths, but normative ideas that guide action. They function as a "regulative ideal" in a distinctively Kantian sense[11], that is even if we were not able to fully determine their ontological status they are still capable to provide normative guidance.[12]

The indirect way of determining the normative relevance of the notion being human is in my view without alternative. It consists in going through an (ideal) process of deliberation and coming to an agreement. This agreement has to be based on the maximum of information available after a process of mutual engagement and explanation of the different opinions held by those

---

[10] I would argue, that even the discussion of such "objective" matters as biological taxonomy is normatively impregnated when talking about the position of human beings. There is something different in classifying funghi or worms or human beings. But more obvious is this special connotation outside of the "purely scientific" context.
[11] See Kant (1787, p. 427 [B 672]).
[12] This claim is informed by a pragmatist approach to ethics, cf. e.g. LaFollette, 2000.

participating in the process. As this process bears similarities with democratic processes, I call it quasi-democratic (cf. Kitcher, 2001).

The participants of this ideal deliberative process[13] must be as numerous as possible. Nobody can be excluded from participation a priori. Everybody conceiving of her or himself as falling under the term human being has to be allowed to participate in this deliberative process.[14] For some individuals not being able to raise their voice themselves – be it because of their age or some handicap – representatives have to be admitted to the process to assure that their opinion is also present in the deliberation.

The participants will be numerous, so one should imagine a gathering of representatives entering a discussion in which they share the information they have with one another. For this they must make themselves understood, even if it is a highly specialised knowledge they want to bring into the deliberative process.[15] This conversation leads to mutual engagement of the participants.

One important condition of this deliberation consists in that it is public. Publicity prevents hidden interests or strategic lies from entering into the ideal process.

Ultimately the result of such an ideal process would consist in a consensus about the normative self-understanding of human beings in form of significant human traits that are considered particularly valuable. This consensus will be quite minimal and it will not cover all particular opinions about valuable human traits. Nevertheless, a small core of valuable significant human traits may emerge as a justified centre of stronger, that is methodologically sound, anthropological arguments. This view could be called "pragmatic essentialism".

The results of such an deliberative process would be capable to withstand the three types of criticism mentioned above: The moral evaluations would designate specific traits agreed upon; their ontological status would not be directly based in "natural facts or properties" independent of human action,

---

[13] My suggestion for an ideal deliberative process takes up ideas from – among others – John Rawls (1971), Jürgen Habermas (1983), and John Dewey (1938). For more details, see Heilinger (2010, part IV).

[14] This condition intentionally allows that some other beings – maybe cyborgs or extraterrestrials – could participate in this process, if they can convincingly claim to understand themselves as human beings. This, of course, is a fictitious assumption, but a necessary condition to avoid any overly simplistic naturalistic understanding of the normativity of being human.

[15] Kitcher has developed a model how to include expertise in democratic decision making processes, Kitcher (2001, part II).

but in a procedure that can be understood by all (who potentially even contribute to this process).

Furthermore its evaluation would not risk to praise obviously morally problematic entities as murder or rape, as the agreement would allow for specific evaluations. And if this procedure were to lead to a narrow and specific selection of certain traits, it would not be problematic, because its selection were based on considered agreement, not on a contingent and arbitrary selection of some natural facts.

And – with regard to the pragmatic objection – history has not shown that considered agreement about the moral status of all human beings has caused the oppression or exploitation of some.

Yet, there are two obvious objections against the suggestion of using an ideal deliberative process to determine the normative content of the notion being human (or to determine the relevance of stronger anthropological arguments). First, it can be doubted that such a process would ever be possible: can such ideal conditions ever be realised? And, second, even if it were possible, would there ever be agreement on the normatively significant traits of human beings?

As I said before, I see the engagement in this deliberative process being without alternatives. If we are not willing to stop the project of determining the normative content of what it means to be a human being altogether, we have to engage in this process. Surely, it will not be providing simple solutions to age old questions and settle ancient disputes within short time. But there is hope that at least some progress can be made in reaching a better understanding of the justified normativity in the notion being human. That is why one would have to try to realise the ideal process even if it should be possible only in a sub-ideal way. The sub-ideal way could consist in many different activities, some of them already going on. Political, societal, international debate about what are the basic elements of human nature that we want to protect from changes are certainly an obvious attempt to engage into the deliberative process. But also more specific debates as in scientific discourse, or in philosophy, sociology etc., play an important role. Furthermore, dialogue between different religions could strive to identify certain common assumptions about what it means to be human. There are already some discussions taking place. But, are there any agreements visible?

Even if a conclusive and stable agreement on what it normatively means to be a human being is out of reach, there might be relatively stable agreements

about basic assumptions. But as there are still new findings about both the human organism itself and its co-existence with others, and as there are also "real" changes of what human beings are and how they live (through cultural evolutions and maybe also through the application of human enhancement interventions), the quest for an "eternal truth" in this matter would be unrealistic. Human beings and their knowledge about themselves are constantly evolving. Consequently, the debate about the normativity of what it means to be a human being must be an open and opening debate, not a closing one.

Still, some results have to be fixed, even if only as preliminary results, if the method suggested here can aspire to be functional at all. To illustrate the fragile agreements (maybe in the form of a reflective equilibrium) and the minimal overlapping consensus that might be found at a given time, I want to use a metaphor, namely speaking of a map of the notion human being which is to be drawn. A map provides orientation according to the needs of those for whom the map was drawn. Even though maps can be "true" to their environment, maps themselves change over time in order to accommodate new demands without that change renders the older maps less "true".[16]

2.3. Four Significant Components of Being Humans

The discourse about what is a human being or what it means to be a human being is rich and multifaceted. Contributions to it stem from different fields (science, philosophy, religion, individual insights etc.) and refer to different aspects of being human (having specific experiences, being a physical organism etc.). In the following I take up the challenge to make an initial suggestion of what might be the result of the ideal deliberate process I have outlined above; i.e. what the content of human self-understanding consists in. The result would be – to stick to the metaphor introduced above – a significant map of the term human being.

Of course my suggestion cannot be other than preliminary: It strives for the description of a fragile, minimal consensus that immediately calls for further examination. It hopefully will stimulate critique and suggestions of alterations and improvements. But, as mentioned above, the deliberative process about what it means to be a human being is an open one, not a closed one. Hence any

---

[16] The metaphor of a map in order to illustrate the claims of "modest realism" has been used by Philip Kitcher in 2001, ch. 5.

contribution or suggested change is welcome and in line with the theory I try to develop. A map is not drawn for ever, but constantly in need of improvements; especially when one gets to know more about the terrain, or if the needs of those using the map alter.

As a single author I cannot perform the deliberative process alone. Hence I suggest another way of putting flesh to the bones of my normative theory. I will draw – yet in a non-systematic and non-comprehensive way – from different sources: From conceptual analysis of the actual use of the notion human being in normative contexts, as well as from the extensive writings within the tradition of philosophical anthropology, philosophy of person, philosophy of mind, biology (currently the developments of neurobiology call for special attention), but also from literature and religious writings. As I said, initially any contribution which can be explained to and scrutinised by the other participants of the debate, can rightly claim attention and is worth to be considered.

In the following I suggest to identify four normatively relevant core components in a "map" of the term human being. These components design significant aspects of what it means to be a human being. The four components vary in scope. But obviously the question what it means to be a human being[17] is a rich question that cannot be answered by reference to a single level of explanation alone.

C1: Human beings are living organisms

Human beings are living beings. This insight might appear to be trivial, but it is both basic and significant for an anthropological mapping of the term human being. Being alive means having a material organism that is characterised by the properties of living beings: e.g. metabolism, growth, reproduction. I do not want to go into the details of the definition of life, but with this first general qualification of human beings it becomes clear that human beings are part of the living world, and with the fact of being alive comes the inevitability of dying.

---

[17] I mean here the basic anthropological question what it means to be a human being in comparison to other entities in the universe, as different from e.g. the taxonomical question regarding the DNA of homo sapiens in comparison to other living beings.

It is significant to conceive of human beings as living beings, because their being alive is the condition for all the other traits and activities that distinguish them.

C2: Human beings have a specifically embodied and embedded mind

As important as the first component of the term human being may be, it is not narrow enough to specifically determine human traits. But the human way of being alive is depending on a particular organic structure: There is a specific human body which can be described in detail. This specific human organism gives rise to a specific form of embodied experience and cognition. Human beings have both a material form and an internal dimension of experiencing. This dualism of aspects of human beings is not meant to constitute two unrelated spheres of being human, quite the contrary. That one can conceive of two different aspects presupposes the mutual conditioning of the physical and phenomenal: Without a physical body we would not have any experiences of the world and of ourselves, and without our experiences we would not have awareness of or interest in us being material bodies.

Anthropologically important is the specific form of human embodiment, which gives rise to a specific form of human cognition and experience. As different human beings share the basic forms of this embodiment, they are capable of sharing experiences. This allows human beings in a basic way to jointly refer to entities and to show empathy. To put it briefly: Because human beings are in their specific embodiment very similar to one another, they are able to share a world of things and experiences and to understand one another.

It is the specific form of shared embodiment that avoids the isolation of human individuals an instead gives rise to a commonly shared realm of experiences and things in which humans can interact with one another.

C3: Human beings are in need for orientation

Humans have – as specifically embodied living organisms – the capacity for spontaneous action. This means that they are not completely determined in their behaviour by instincts or by some hard-wired brain or gene structures, but that they have alternative possibilities to live their lives. These different possibilities cause human beings to require orientation. As they can choose to do one thing or another humans are looking for guidance. This guidance is

most often primarily provided by peers, by the social or cultural environment that influences the perceived options to act and to think.

One important form of orientation is given by the cultural self-understanding of human beings. Anthropological thinking is itself the explicit questioning of the implicitly action-guiding background assumptions of what human beings are and what human actions should be like.[18]

There is more than one way to conceive of human beings; hence the ongoing competition between the different interpretations of human beings. For example, currently a dispute is taking place between religious conceptions of being human on the one side, and political conceptions of being human on the other side. Furthermore there is a conflict between the scientific interpretations of human beings – as stressing determinism by brains and genes – versus the self-understanding as free and responsible agents on the basis of the individually perceived possibility to choose between different behaviours.

The need for orientation in face of different possible actions and different ways to conceive of human beings is a significant component of the term human being. If it were not for this basic openness and possibility to orient themselves (for better or for worse), human beings would be running a programme or living randomly, and in either case not be responsible for their actions.

C4: Human beings are "anthroponomous"

The fourth significant component of the term human being is based on the other three components mentioned above. Human beings are – as living beings with a specific body and specific mental capacities that allow them to be not completely determined by physical facts – able to decide autonomously about their actions, and to interpret and define themselves. This capacity for autonomous decision making does not stand in opposition to natural facts, but rather takes place under the condition of humans being material beings.

What actions human beings perform and which self-descriptions they accept, depends to a high degree on themselves. Human beings are self-

---

[18] As human beings need this kind of orientation through certain assumptions about what they are, one could call them "menschenbilderbedürftig", they stand in need for cultural images about themselves. Cf. Müller and Heilinger, 2008.

interpreting and because of this also self-determining beings.[19] Such a human self-interpretation and self-determination is not completely free floating but stands under the conditions of them being natural beings. Such interpretation and determination are, as it were, part of their nature. But there are different ways to concretely fill out the multiple possibilities human beings have to lead their lives or to conceive of themselves.

If one were looking for a name for this capacity, I would suggest to call it "anthroponomy", a compound from greek anthropos (human being), and nomos (law, order, custom, determination, definition). The compound "law of humans" or "humans' law" could be understood in two ways. First, as an objective genitive, which specifies that there is some definition or determination of human beings; second, and most important, as a genitive subjective construction, indicating the origin of this determination which governs human behaviour and human self-understanding: it stems from human beings themselves.

The four components are suggestions of what could qualify as significant elements in the term human being. It is not meant to provide an exhaustive description of what human beings are, but a minimal designation of important facts about humans. I claim that these four components might be agreed upon by potentially all human beings as apt descriptions of what human beings are.

Of course, this is nothing but a very minimal description and there is much more to be said about human beings. But the challenge was to find significant components of the notion being human, to which most – if not all – could agree. Even if the result is only a small overlapping consensus, it is still of use. It might support the view that anthropological arguments should only focus on a restricted set of propositions about human beings.

2.4. An Example: Life-Extension

To employ these arguments with regard to concrete enhancement interventions would demand more extensive discussion than I can provide here, so a short illustrative example has to suffice. Imagine it would become possible for human beings to radically improve their healthy and active life-span. For the sake of the anthropological argument, it should be assumed here

---

[19] For this, see, e.g., Charles Taylor's work on philosophical anthropology (Taylor, 1971 and 1985).

that the risks involved in this intervention are minimal, that there has been found a way to deal with the societal challenges coming along with a radically prolonged life for example with regard to pension payments and that everyone who underwent this intervention has given her fully informed consent. Of course it is highly improbably that these conditions will ever be met, but my question here is: How would anthropological arguments assess such an intervention if no other moral considerations would speak against it?

For answering this question it may be interesting to keep in mind that in the Western world the average life expectancy continuously increased in the past. This has so far not brought about any danger for our self-understanding as human beings. Humans still are living, interacting beings, with a capacity to autonomously make the necessary choices in life. So it seems reasonable to assume that a stepwise further increase of the average life expectancy can be coped with without endangering human nature. But imagine, humans would now be able to radically expand their life-expectancy from say about 80 to about 250 or 500 years, or even longer. Would that have a detrimental impact on the self-understanding of humans as human beings? My suggestion was to answer this question with regard to the four normative components[20] that determine what it could mean to be a human being. Obviously, as long as humans do not become immortal, the first significant element of human nature remains intact. Equally, there seems to be no reason to doubt that humans will remain embodied, sentient and interacting beings that are not fully determined in their choices and hence stand in need for orientation. So the second and third significant components would remain unchanged even if human lives would suddenly last much longer. It is only with regard to the fourth component of being human that anthropological arguments can raise some moral objections about radically extending the healthy human life span. For this, think of an argument provided by Bernard Williams in his discussion of the "Makropoulos case" (Williams, 1973). Williams argued that a radical increase in the amount of time available to human agents would decrease the necessity and urgency to make choices and to act here and now since it would always be possible to postpone action to a later time. He illustrates his claim by referring to "E.M.", the long-living heroine in Leoš Janáček's Opera Věc Makropulos. While a stepwise increase of life-expectancy that does not goes

---

[20] As argued above, these components are tentative and open for revision by the quasi-democratic deliberative process.

beyond a certain threshold could match with the human intellectual set-up to act autonomously and make life-plans, a radical increase beyond this threshold would possibly endanger agency altogether. Autonomy might be lost, if the purpose of acting and living fades out by exceeding a manageable and rather short life span.

Obviously, further argument would be necessary to determine more precisely the permitted pace of increasing the average life-expectancy and also the threshold beyond which making life plans that presuppose autonomous actions here and now. Yet, the example should illustrate *that* anthropological arguments even in their minimal form as defended here *can* indeed come up with constraints against enhancement interventions. The anthropological constraints about radically increasing the healthy human life span – in the hypothetical absence of any other moral constraints based on considerations of risk, justice or doubts about informed consent to the intervention – are admittedly very basic. This shows that anthropological arguments are best understood as widely permissive and not as restrictive as often suggested by proponents of "bioconservativism". Indeed, human nature seems capable to integrate much change and to accommodate diverse forms of human life. Parochial thinking about what is familiar should not be taken for moral arguments against possible change. Biological and also cultural evolution have shown that change is essential to life, and also to human life.

Conclusion

In this paper I discussed the relevance of anthropological arguments, exemplarily in ethical debate about human enhancement interventions that aim at altering human traits or functioning or at pushing the boundaries of human nature. I have argued that anthropological arguments can be justified and specified in their content with the help of an idealised process of "quasi-democratic deliberation". However, such anthropological arguments stand not alone in evaluating these biotechnological interventions; there are also justice-related, risk-related, and autonomy-related issues.[21] From the point of view of anthropological arguments, the evaluation of human enhancement interventions tends to be most often permissive, since the justifiable content of

---

[21] There might be convergence between the different types of ethical judgements. However this classification is not meant to establish firm boundaries but rather to give an orientation about the different approaches to assessing the ethical challenge at hand.

anthropological arguments is rather thin: Often, they do not support seeing a "threat for human nature" or a potential de-humanisation in human enhancement interventions. It is only in some rather extreme cases that anthropological arguments would be able to speak against certain interventions. The main part of the debate focusses correctly on the more pertinent dimensions of justice, autonomy, and risks. But still, anthropological arguments are elementary. They play an important, basic role in the background of the debate, insofar as they discuss and make explicit the fundamental orientations about what it means to us to be human beings. Furthermore, they are elementary, because elements of this basic debate find their ways also in the other layers of the debate: risks are risks for human beings, because they threaten to harm the specific human way of living and well-being; justice matters, because from an anthropologically informed perspective, we judge every human being to be basically equal in value and moral standing; and autonomy is of relevance, because it crucially matters to human beings that they can either have it or not.

A last clarification: The anthropological arguments alone are incapable of providing substantive reasons not to engage in interventions aiming at overcoming the human condition and radically transgressing human boundaries. But they show, that if we did, we would enter a post- or non-human stage in which anthropological arguments would have lost their bite simply because they would not apply any more. Yet, as long as we stay in the human realm, sound elementary anthropological arguments do provide fundamental moral orientation. In particular, they call for moderate changes that can be caught up with in a deliberative process of collective self-determination as human.

REFERENCES

Beauchamp, T.L., & Childress, J.F. (2009). *Principles of Biomedical Ethics.* 6th edition. New York/Oxford: Oxford University Press.

Birnbacher, D. (2006). *Natürlichkeit.* Berlin/New York: de Gruyter.

Bostrom, N. (2003). Human Genetic Enhancements: A Transhumanist Perspective. *Journal of Value Inquiry, 37*(4), 493–504.

Bostrom, N. (2008). Why I Want to be a Posthuman when I Grow Up. In B. Gordijn, & R. Chadwick (Eds.), *Medical Enhancement and Posthumanity*, Berlin: Springer Netherlands, 107–136.

Buchanan, A. (2009). Human Nature and Enhancement. *Bioethics, 23* (3), 141–150.

Dewey, J. (1938). *Logic. The theory of inquiry* (The later works vol. 12, ed. by Jo Ann Boydston). Carbondale: Southern Illinois University Press 1986.

Fischer, J. (2008). *Philosophische Anthropologie. Eine Denkrichtung des 20. Jahrhunderts.* München: Karl Alber.

Habermas, J. (1983). Diskursethik: Notizen zu einem Begründungsprogramm. In J. Habermas (Ed.), *Moralbewußtsein und kommunikatives Handeln*, Frankfurt: Suhrkamp, 53–125.

Heilinger, J.-C. (2010). *Anthropologie und Ethik des Enhancements.* New York: de Gruyter.

Heilinger, J.-C., & Crone, K. (2014). Human freedom and enhancement. *Medicine, Health Care, and Philosophy 17 (1),* 13–21.

Kant, I. (1787). *Kritik der reinen Vernunft* (2nd edition [B]). In Kants gesammelte Schriften, Königlich Preußische Akademie der Wissenschaften (ed.), Volume 3, Berlin (1904/11).

Kitcher, P. (2001). *Science, Truth, and Democracy.* Oxford/New York: Oxford University Press.

LaFollette, H. (2000). Pragmatic Ethics. In H. LaFollette (Ed.), *The Blackwell Guide to Ethical Theory*, Oxford: Blackwell, 400–419.

McConnell, T. (2010). Genetic Enhancement, Human Nature, and Rights. *Journal of Medicine and Philosophy 35*, 415–428.

Mill, J.S. (1874). Nature. In J.S. Mill (Ed.), *Essays on Ethics, Religion and Society*, Indianapolis 2006: Liberty Fund, 373–402.

Moore, G.E. (1903). *Principia Ethica.* Cambridge: Cambridge University Press.

Müller, O., & Heilinger, J.-C. (2008). Zehn Thesen zur "Natur des Menschen". Grundriss zu einer Anthropologie. In D. Ganten, V. Gerhardt, J.-C. Heilinger

and J. Nida-Rümelin (Eds.), *Was ist der Mensch?*, Berlin/New York: de Gruyter, 190–196.

The President's Council on Bioethics (2003). *Beyond Therapy. Biotechnology and the Pursuit of Happiness.* New York: Dana Press.

Rawls, J. (1971). *A Theory of Justice.* Cambrigde, MA: Harvard University Press.

Roduit, J., Baumann, H., & Heilinger, J.-C. (in press). Human Enhancement and Perfection. *Journal of Medical Ethics.*

Sandel, M. (2007). *The Case Against Perfection. Ethics in the Age of* Genetic Engineering. London: Harvard University Press.

Taylor, C. (1971). Interpretation and the Sciences of Man. *Review of Metaphysics, 25*, 3–51.

Taylor, C. (1985). Self-interpreting animals. In C. Taylor (Ed.), *Philosophical Papers 1*, Cambridge: Cambridge University Press, 45–76.

Williams, B. (1973). The Makropoulos Case. Reflections on the Tedium of Immortality. In B. Williams (Ed.), *Problems of the Self,* Cambridge: Cambridge University Press, 82–100.

Identity Expansion and Transcendence[*]

William Sims Bainbridge[†]
wbainbri@nsf.gov

ABSTRACT

Emerging developments in communications and computing technology may transform the nature of human identity, in the process rendering obsolete the traditional philosophical and scientific frameworks for understanding the nature of individuals and groups. Progress toward an evaluation of this possibility and an appropriate conceptual basis for analyzing it may be derived from two very different but ultimately connected social movements that promote this radical change. One is the governmentally supported exploration of Converging Technologies, based in the unification of nanoscience, biology, information science and cognitive science (NBIC). The other is the Transhumanist movement, which has been criticized as excessively radical yet is primarily conducted as a dignified intellectual discussion within a new school of philosophy about human enhancement. Together, NBIC and Transhumanism suggest the immense transformative power of today's technologies, through which individuals may explore multiple identities by means of online avatars, semi-autonomous intelligent agents, and other identity expansions.

Introduction

For over three centuries, the forces unleashed by the Enlightenment and the Industrial Revolution have been eroding traditional notions of human nature, even as they have been expanding the powers of human creativity. In the present century, we may have reached the point at which it may be impossible to say what we *are*, even as we can decide what we *will be*. Two parallel

---

[*] The views expressed in this essay do not necessarily represent the views of the National Science Foundation or the United States.
[†] National Science Foundation, USA.

intellectual movements, Transhumanism and Converging Technologies have been developing rather deep conceptualizations of this dynamic situation, largely independent of each other, and one theme of this essay will be how these two historical strands can be knit together. Two specific issues will anchor the discussion, personality expansion and death transcendence, both to a significant extent achievable through the application of advanced information technology.

Since the first time a proto-human intentionally chipped a piece of stone to make a cutting tool, technology has extended the scope of human action, but now such developments as artificial intelligence permit extension of our very identities. Fear of one's own death, and grief over the death of loved ones, largely motivated the invention of religion, which is severely threatened by modernity. Advanced technology offers more than merely transcendence of death, but implies a reformulation of human nature and thus of societal institutions like churches. These issues can best be addressed after considering the origins and histories of the two intellectual movements.

1. Transhumanism

Primarily philosophical and cultural, Transhumanism has many roots and branches, so it is impossible to set an exact boundary on this movement. An obvious core is a small network of friends, primarily in the United States and Britain but also including people in many other nations, who communicate intensively with each other and have held various conferences and set up a dynamic series of formal organizations with names like The Extropy Institute, Terasem, The World Transhumanist Association, and Humanity+ (H+). In addition to the somewhat organized core of the movement, a significant number of very famous individuals have charted their own transhuman courses and intermittently interact with each other and with the organizations. Each member, in his or her own way, draws upon older intellectual traditions. One way to get an efficient overview is to scan the online self-descriptions of four of the related groups:

> Humanity+ (humanityplus.org and hplusmagazine.com):
> ...dedicated to elevating the human condition. We aim to deeply influence a new generation of thinkers who dare to envision humanity's next steps. Our programs combine unique insights into the developments of emerging and speculative technologies that focus on the well-being of our species and the changes that we are and will be facing. Our programs are designed to produce outcomes that can be helpful to individuals and institutions. Humanity+ is an

international nonprofit membership organization which advocates the ethical use of technology to expand human capacities. In other words, we want people to be better than well.

Institute for Ethics and Emerging Technologies (www.ieet.org):
...formed to study and debate vital questions such as: Which technologies, especially new ones, are likely to have the greatest impact on human beings and human societies in the 21st century?   What ethical issues do those technologies and their applications raise for humans, our civilization, and our world? How much can we extrapolate from the past and how much accelerating change should we anticipate?   What sort of policy positions can be recommended to promote the best possible outcomes for individuals and societies?

Terasem Movement, Inc. (www.terasemcentral.org):
...a 501c3 not-for-profit charity endowed for the purpose of educating the public on the practicality and necessity of greatly extending human life, consistent with diversity and unity, via geoethical nanotechnology and personal cyberconsciousness, concentrating in particular on facilitating revivals from biostasis. The Movement focuses on preserving, evoking, reviving and downloading human consciousness.

Singularity University (singularityu.org):
...a unique interdisciplinary, international and intercultural experience which challenges students to use transformative, exponential technologies to address grand challenges. SU educates and inspires students to discover and design sustainable organizations to positively impact humanity.

Transhumanists frequently cite Friedrich Nietzsche's concept of the Übermensch, presented well over a century ago in somewhat obscure form in the philosopher's quasi-biblical masterwork *Also Sprach Zarathustra* (Nietzsche, 1885; Sorgner, 2009; Bainbridge, 2010). The Übermensch is a superior form of being, toward which humans should aspire, although its exact nature and the route to its accomplishment may be beyond our present comprehension. A primary question for Nietzsche himself was how individualistic versus communal the Übermensch should be, a debate outlined in his earlier book *The Birth of Tragedy*, and this remains an issue (Nietzsche, 1872). In the first half of the twentieth century, many science fiction authors wrote often rather deep novels about superior humans, which were the product of natural evolution, mind-training systems, or advanced technologies.

Examples include *Odd John* by Olaf Stapledon (1935), *Slan* by A. E. Van Vogt (1946), and *More Than Human* by Theodore Sturgeon (1953).

Given the individualism and powerful achievement motivations of many current Transhumanist leaders, it is difficult to assemble an authoritative history of the early days of the movement, because each leader has a different recollection of it including the role they themselves played. Clearly, a key individual was Fereidoun M. Esfandiary, son of an Iranian diplomat who settled in the United States and changed his name to FM-2030 to symbolize both the rejection of constraining ethnicity and the embrace of future possibilities. His 1989 book, *Are You a Transhuman?*, not only set forth many principles of what became H+ but was organized around an extensive questionnaire that could be used to document the values and personality of the reader preparatory to various transformation methodologies (FM-2030, 1989).

Recently, two of FM-2030's associates, Max More and Natasha Vita-More (2013), published a massive edited volume nicely summing up this heritage, *The Transhumanist Reader*. Ten of the 42 chapters were reprinted from periodicals they had edited in the period 1992-2002 under the name *Extropy*, a term More defines as "the extent of a living or organizational system's intelligence, functional order, vitality, and capacity and drive for improvement" (More, 2013, p. 5). Transhumanists sometimes use the term *extropy* as an antonym for *entropy*. They generally do not use the alternative term from physics, *enthalpy*, nor *élan vital* which Henri Bergson (1911) considered to be the impetus responsible for evolution, but these are similar concepts. Another 19 chapters were newly written expressly for the *Reader*, and the remaining 13 were reprints from a variety of sources. Several of the contributors are quite prominent intellectuals, notably the artificial intelligence pioneer Marvin Minsky, who brought copies of his book *The Emotion Machine* (Minsky, 2006) to one Terasem meeting I attended, and Singularity University futurist Ray Kurzweil who discussed his book *The Age of Spiritual Machines* (Kurzweil, 1999) with me at another.

The nine sections of the *Reader* offer an ontology of the movement, beginning logically enough with an exposition of philosophical themes, including an eight-point Transhumanist Declaration that ends by proclaiming a new human right: "We favor morphological freedom - the right to modify and enhance one's body, cognition, and emotions. This freedom includes the right to use or not use techniques and technologies to extend life, preserve the self through cryonics, uploading, and other means, and to choose further

modifications and enhancements" (More, Vita-More, 2013, p. 55), *Cryonics* is the practice of freezing the entire body or just the head of a recently deceased person, in hopes that future biological technologies will be capable of restoring the individual to life and health. FM-2030 is currently under cryonic suspension, and Max More is the president of the Alcor Life Extension Foundation, the most prominent organization providing this service. *Uploading* refers to the transfer of human personalities into computerized information systems, hopefully allowing them to continue to function and interact with living people after their original biological substrates have ceased to function. That is the area within Transhumanism where I myself have been most active. My new book, *Personality Capture and Emulation*, describes the current state of the technology and sets out an agenda for further development (Bainbridge, 2014). My *Reader* essay and my other recent book, *eGods*, outline how people can cooperate to achieve cybernetic immortality, through avatars of a deceased person under the control of a living person (Bainbridge, 2013a, 2013b).

The second and third sections of the *Reader* cover human enhancement in the somatic and cognitive spheres, and the fifth section explicitly concerns death transcendence. Other sections describe the core technologies, consider ethical and political implications, and offer the optimistic view that technological development is accelerating toward a singularity at which almost anything will be possible. The final section responds to the assertion by conservative intellectual Francis Fukuyama (2002, 2004) that Transhumanism is the most dangerous idea in the world, worse even than Islamic radicalism, thus deserving to be suppressed.

A more moderate debate , roughly parallel to the cryonic-upload and somatic-cognitive distinctions, concerns whether the most promising technologies for human modification are biological or computational. Brian Wowk's (2013) contribution to the Reader's section on death transcendence takes the biological approach, and was reprinted from *The Scientific Conquest of Death: Essays on Infinite Lifespans*. I also had contributed to that earlier anthology, with an essay on computational personality capture (Bainbridge, 2004), that was published right between similar essays contributed by Ray Kurzweil (2004) and Marvin Minsky (2004). A later collection with the same perspective, *Unnatural Selection: The Challenges of Engineering Tomorrow's People*, included chapters by several *Reader* contributors, among them another of my personality capture studies (Bainbridge, 2009).

Whatever disciplines they earned their academic degrees in, most Transhumanist leaders function as philosophers, proposing and analyzing ideas and ideals, rather than conducting empirical research or constructing technologies. While imaginative, members of the Converging Technologies movement exhibit the opposite pattern, anchoring their work in practical accomplishments. Thus the two parallel movements have achieved an efficient division of labor between them, and those of us who belong to both believe they are partners creating the future for all humanity.

2. Converging Technologies

Primarily a technologically sophisticated community for research and engineering, Converging Technologies grew out of an intellectual movement initially focused on nanoscience and nanotechnology, which subsequently added biotechnology, information technology and new technologies based on cognitive science to go through a decade-long phase when it was called NBIC for Nano-Bio-Info-Cogno. The primary mode of organization and communication has been major conferences sponsored by US government science agencies, often with international components, that brought together experts across many fields from academia, industry, and government. Book-length reports always resulted from these conferences, often influential in shaping support for research around the globe (Roco & Bainbridge, 2003; Roco & Montemagno , 2004, Bainbridge & Roco, 2006a, 2006b). The latest such report was based on workshops held around the world, and organized from the US National Science Foundation (Roco et al., 2014). Participants were generally leaders in conventional institutions, and included only a few people directly associated with the Transhumanist social movement. Five of the contributors to *The Transhumanist Reader* also contributed to the Converging Technologies reports, in addition to myself: Wrye Sententia (2004, 2006), James Hughes (2006), Andy Miah (2006), and Anders Sandberg and Nick Bostrom (2006).

The historical origins of nanotechnology are open to debate, because two competing models can be argued. First, research and engineering at the nanoscale - structures less than 100 nanometers in size but larger than single atoms - was a natural extension of normal-science work in physics, chemistry, materials science, and microbiology. Second, it may have resulted from a revolutionary-science movement because its public face in the 1990s was

science fiction inspired by non-technical publications written by Eric Drexler, who is a technically sophisticated visionary, but not exactly a scientist or engineer. I prefer to look at it in a third way, more oriented toward the future than the past, but beginning with a conceptualization of prior human evolution (Bainbridge, 1997).

The universe was not created by a god, but selected by a human being, namely yourself. Of course, this is a metaphor, but worthy of consideration as a step toward a more formal model that might be developed in the future. There already exists a philosophical idea known as the Anthropic Cosmological Principle (Carr & Rees, 1979, Gale, 1981; Leslie, 1982; Barrow & Tipler, 1986), opposed to the more traditional Argument from Design.

Why is the world conducive for human life? Because God made it so. That is the Argument from Design: God must exist because there is no other explanation for the benevolence of nature toward humans, than that some superior being created the world with that purpose in mind (Bertocci, 1945). Corollaries of this theory are that the universe is meaningful, that its meaning centers on human beings, and that humans can trust that the future will be good. A century ago, when science had begun to discover how complex nature really was, biochemist and sociologist Lawrence Henderson (1913, 1917) pondered the "fitness of the environment" in a way that suggested science could be itself part of God's design, but admitting that our existence remained something of a puzzle. Today's scientists and engineers may not be religious, or contemplate the Argument from Design, but they seem committed to the view that endless progress is indeed possible, which would be the case if it were central to God's plan.

The Anthropic Cosmological Principle offers an alternate explanation, atheistic and less optimistic, in which the word *anthropic* places humans rather than gods at the center of the mystery. Here is one variant of it. The universe is vast and complex, containing a great variety of environments, possibly infinite in its diversity. At one location in space and time, the planet we call Earth formed at the proper distance from a G-type star to possess oceans, a non-toxic atmosphere, and other conditions required for the emergence of life. The process of emergence may have involved very low probability local events assembling the first self-reproducing molecules, but then natural selection from random variation led to evolution toward ever more complex systems including intelligence that approached what I have called the *Omicron Point*. This term is by analogy with the *Omega Point* postulated by Pierre Teihard de

Chardin (1964), defining the future goal that God has set for evolution. Both omicron and omega are letters of the Greek alphabet translated into ″O″ in the Latin alphabet, but omicron comes near the middle in alphabetical order, suggesting that the key point in history is not its conclusion, but an intermediate fulcrum around which history pivots.

Omicron is the point in history when intelligence first seriously asked the question of why the universe is conducive to its existence, with the cognitive capability of understanding ideas like the Argument from Design or the Anthropic Cosmological Principle. Unfortunately, if this theory is in fact correct, progress after the Omicron Point is not merely uncertain but unlikely. This is true because the specific time and place when and where omicron occurs is the result of random factors, and one consequence is that the coherence of the environment will degrade over time, but perhaps slowly. Socio-cultural conditions conducive to progress may degrade more quickly. But perhaps most crucial is that fact that natural selection from random chaos does not assure that any additional scientific discoveries or technological inventions will be possible, in addition to those required to reach omicron.

An example seemingly remote from identity expansion, but illustrating key principles, is the currently stalled state of human spaceflight. We simply do not possess the launch technology required for colonization of the solar system, and the other planets are so different from Earth that they would not be economical locations for colonies, an especially obvious point when we remember that Antarctica has research bases but no ordinary towns. Already in 1961, the upper stages of the Saturn I launch vehicle used the most energetic practical rocket propellants, liquid oxygen and liquid hydrogen. Serious research on nuclear engines for launch to orbit was ended in 1972, the same year as the last expedition to the moon, and environmental concerns would prevent any resurrection of this higher-energy propulsion method. Cancellation of the Space Shuttle program indicates that innovative engineering designs really cannot find successful ways to work around these natural laws.

The severe natural limitations on human space travel can be understood in the light of the Anthropic Principle. Precisely because extreme diversity of environments is a requirement for one of them to support life, a solar system is exceedingly unlikely to possess more than one habitable planet. For there to be two temperate Earths, they would need to be in similar orbits, which over time would have lead to gravitation-induced orbital instabilities, either causing

the two to collide, or moving them into very different orbits, both of which could not be at the right distance from the Sun to support liquid water and thus life. If the Earth were smaller and thus had weaker gravity, it would be easier to launch spacecraft from it, but the planet would be unable to hold a rich atmosphere for the hundreds of millions of years required for the evolution of intelligent life. Travel to other stars that might have habitable planets would be easier if they were closer, but then the chances would have increased that Earth would have been destroyed when our solar system collided with another. While there is much room for debate about the potential future of astronautics, these points are easy to understand, and they stress the difficulties science and technology will have more generally, sustaining progress far after the Omicron Point.

This returns us to a consideration of Converging Technologies, which can be seen as a vigorous attempt to keep progress moving forward, even in the absence of divine aid. Whether or not Eric Drexler's ideas were really influential in the development of nanotechnology, he is a good example of the issue. Earlier in his career he was highly active in the L-5 society, a visionary but technically competent movement to advance human settlement of outer space. The name L-5 came from the idea of building an orbiting city at the Lagrange 5 point in the Moon's orbit. Founded only 3 years after the last human expedition to the Moon, at first L-5's goal may have seemed feasible to many reasonable people, but hope does not really "spring eternal." In the 1980s, L-5 merged with the National Space Institute to become the National Space Society, and while some local groups continued to uphold the L-5 vision, this moderating merger reflected the inability of humans to undertake colonization of the solar system. Drexler shifted his creative energies to writing books about the potential for progress at the opposite end of the distance scale from the vastness of outer space, to the infinitesimal inner space where nanotechnology operates.

The 1959 lecture "Plenty of Room at the Bottom" by physicist Richard Feynman influenced Drexler, who was not a participant in the Converging Technologies conferences, and it may have inspired many participants as well (Bainbridge, 2007, pp. 35-41). The implicit meaning of Feynman's title concerns where science may find scope for new discoveries, new intellectual or physical territories beyond the current frontier of knowledge. Feynman listed a number of radical technological innovations that might be achieved on very small scales, including but not limited to the nanoscale which is conventionally defined as between 1 and 100 nanometers on at least one of its dimensions.

Once the US National Nanotechnology Initiative was established in 2000, two processes expanded its scope to create a general progress-oriented movement.

First, actual nanotechnology developments tended to be in three areas that connected to other fields. Feynman and Drexler had emphasized the idea of building nanoscale machines following the principles of large-scale mechanical devices, but relatively little progress was achieved in this subfield, both because it proved very difficult to form and assemble their parts, and because natural phenomena at the lower end of the nanoscale, notably van der Waals forces, operate against the functioning of wheels and gears. Thus pure nanotechnology tended to become a subfield of materials science, for example through research on nanoscale particles similar to fine powders and on thin layers of one material on top of another, such as the coating of computer disk drives that utilized the nanoscale principle called the *giant magneto resistance effect* to store more data in small areas. A second nano-relevant field was microbiology, because proteins and DNA consist of molecules of nanoscale dimensions. The third was information technology, especially when transistors and other electronic elements on solid-state chips were successfully engineered with dimensions less than 100 nanometers. Thus, around the year 2000 it was quite natural for experts in the field to think about the synergies that could be achieved by bringing together nanotechnology, biotechnology, and information technology, nano-bio-info or NBI as it could be called.

Second, in part because of the connection to biotechnology and potential medical applications of nanotechnology itself, great attention was given to the social and ethical implications of nanotechnology (Roco & Bainbridge, 2001, 2006a, 2006b). The issues could be as straightforward as the danger of release of large quantities of nanoparticles from industry into the environment, or as complex as the second-order unintended consequences of new human-centered nanotechnologies. It is worth noting that computer and information science has given less emphasis to social and ethical issues than have medical technologies and nanotechnology, and increased attention to potential harmful effect of information technology would be a logical consequence of its convergence with nano and bio. The website of the National Nanotechnology Initiative asserts:

An important component of responsible development is the consideration of the ethical, legal, and societal implications of nanotechnology. How nanotechnology research and applications are introduced into society; how transparent decisions are; how sensitive and responsive policies are to the

needs and perceptions of the full range of stakeholders; and how ethical, legal, and social issues are addressed will determine public trust and the future of innovation driven by nanotechnology.[1]

The first major NBIC conference, held at the National Science Foundation, produced a book-length report whose title explicitly stressed revolutionary possibilities: *Converging Technologies for Improving Human Performance* (Roco & Bainbridge, 2003, p. 3):

> At this unique moment in the history of technical achievement, improvement of human performance becomes possible. Caught in the grip of social, political, and economic conflicts, the world hovers between optimism and pessimism. NBIC convergence can give us the means to deal successfully with these challenges by substantially enhancing human mental, physical, and social abilities. Better understanding of the human body and development of tools for direct human-machine interaction have opened completely new opportunities. Efforts must center on individual and collective human advancement, in terms of an enlightened conception of human benefit that embraces change while preserving fundamental values.

Such sentiments imply that in future the social sciences would need to be integrated with nano, bio, and info, but already in 2000 it was obvious that a related field needed to be added immediately, namely cognitive science. This relatively new field is itself multidisciplinary, including inputs from neuroscience and artificial intelligence, and thus already connected to bio and info. While the wider NBIC convergence integrated the quartet nano-bio-info-cogno, the triad BIC is also a proper focus of attention, are the dyads bio-cogno and info-cogno. Indeed, the dyads mirror the two specializations usually treated separately in Transhumanism, human enhancement via biotechnology versus via information technology. Given that my own work focuses on the triad info-cogno-social, that is the territory of Converging Technologies I shall emphasize here.

3. Identity Expansion During Life

All forms of life affect their environments, and for tens of thousands of years humans have been creating objects that expressed their thoughts, whether chipping flint to made axe heads of a distinctive style, painting the walls of caves or even their own bodies, and eventually by writing words that can be

---

[1] http://www.nano.gov/you/ethical-legal-issues, accessed September 19, 2013.

understood centuries later. One of the best ancient examples is Julius Caesar, whose written words are still worth reading today, whose face is recognizably preserved in sculptures that basically agree with each other, and through deeds that for better or worse shaped the history of the world. Since the Renaissance, technologies have democratized personality preservation, through the printing press, sound and video recordings, and now the multi-media communications of Internet. Most significantly, modern information technology allows an individual to offload aspects of the self in realtime, not merely preserving them for posterity, but providing much greater scope for action during life. Potentially, each person can become a team, led by the biological person, but uniting multiple semi-autonomous intelligent agents that act as secondary selves cooperating while performing different tasks in parallel.

An appropriate if modest example comes from research I did in 1986 in the wake of the accident that destroyed the Space Shuttle Challenger (Bainbridge, 1991, pp. 75-81). At the moment of the disaster, I happened to be visiting at Jet Propulsion Laboratory, exploring the possibilities for a new research project related to the space program, and watched the catastrophic launch on NASA's direct video feed from Cape Kennedy, along with scientists and journalists who had gathered at JPL for the encounter of the Voyager II Space probe with the planet Uranus. I had done a pilot project developing a questionnaire asking respondents to rate 49 different justifications for the space program, and the Challenger disaster inspired a larger effort that led to a questionnaire in which 894 Harvard students rated 125 potential goals of space exploration. I had used the standard factor analysis method of computerized statistical analysis in the pilot study, to see how the 49 rather specific ideas clustered into larger concepts, each representing a general value that space exploration might serve. However, in 1986 I did not have access to computer software that could do a factor analysis on 125 variables at once, so it was necessary to write my own software, and I decided to do so in a somewhat distinctive manner.

I wrote a set of programs that would allow me to enter the data into my already outdated Apple IIe personal computer, calculate the 15,500 Pearson's *r* correlation coefficients between all pairs of goals, then rearrange the correlation matrix to find blocks of goals that clustered together. The algorithm for the last of these steps was based on work a decade earlier by Harrison White, one of my Harvard professors, called *block modeling* (White et al., 1976). For each of several runs, I would select a threshold criterion,

such as 0.50. Correlations range from -1.00 to +1.00, and with this criterion every correlation in the matrix of 0.50 or greater would be turned into a 1, and all the others into zeros. The program would then go through the same set of steps tens of thousands of times: (1) select two goals at random, (2) calculate X the total distance of all the 1s in the table from the diagonal, (3) imagine switching the position of the two goals in both rows and columns of the table, (4) calculate Y the new total distance of all the 1s from the diagonal, (5) if Y < X, do switch the positions of the two goals, otherwise, do not. This block modeling would go on for as much as 36 hours for each criterion I tried, on an admittedly slow computer even for 1986.

Note that I could have used many different algorithms, for example not replacing the correlation coefficients with 1s and 0s but multiplying each coefficient by its distance from the diagonal. But I had written the program in the way I wanted to, so it was an extension of myself, and even of my personal experience at Jet Propulsion Laboratory and as a student of Harrison White, as well as being an objectively competent research tool. The program and the Apple IIe then served as my semi-autonomous agent, working away according to my instructions, even when I was asleep. They were an extension of me, expanding my ability to do work. Of course, this is not a new principle, as even ancient farmers could rely upon their cows to eat grass and produce milk, even while the farmer himself napped in the shade of a tree. Yet because it is a computer-based example, it illustrates how humans will be able progressively to offload more and more of their work – and of themselves – onto increasingly intelligent machines as the years pass.

A wide range of examples exist today. Trading agents assist investors, from simply issuing stop-loss orders if the value of a security drops some amount specified earlier by the user, to more complex systems for buying and selling. In some massively multiplayer online role-playing games, such as *EVE Online* and *Fallen Earth*, avatars can continue to do work after the player has logged out. People who have rated movies on a recommender system like Netflix are constantly providing automatic advice to other Netflix customers, because the system uses their preferences through complex computations to advise other people who have rated some of the same ones, thereby simulating all the people who rated the movies. Even Google acts as a surrogate for thousands of people, because a main component of its automatic search algorithm is based on links to a given website manually placed by people on their own websites, thus

incorporating the judgments of webmasters whose own sites have found favor with many other webmasters.

Given the diversity of even this very preliminary list of examples, many quite different conceptual approaches can plausibly outline the future possibilities. Here I shall summarize the approach I took in my recent book, *Personality Capture and Emulation* (Bainbridge, 2014). It begins with consideration of two very traditional methods that sought to capture aspects of human personalities for scientific purposes, the culture and personality research carried out by a huge team of scientists at Harvard University over half a century ago, and the emergence of questionnaire-based survey research of public opinions that was evolving at many institutions during the same period of time. Bringing the narrative up to the present time, the book then examines massive questionnaires that may included thousands of questions, which becomes quite practical if they are administered online and through mobile devices, so that it is comfortable for users to answer the questions at odd times throughout the day, over a period of months, for example while waiting for a bus or otherwise lacking anything inherently interesting to do other than to enter one's opinions and feelings into a database.

Three pilot study examples demonstrated feasibility. In one study, 2,000 situations that might trigger one of 20 standard emotions were solicited through an online questionnaire and culled from online publications such as novels. Then a program was written from scratch for a pocket computer to administer 2,000 one-sentence statements describing the situations, which the respondent was asked to rate along a scale for each of the 20 emotions, resulting in fully 40,000 questions, which a test subject had no difficulty finishing during free time over a period of weeks. A second pilot study used an Android app to administer a 200-item standard psychology test online, through a system that provided immediate feedback to the respondents and automatically collated the responses from the more than three thousand people who completed the questionnaire within its first week. The test traditionally was used to measure the Big Five personality dimensions, but given the large number of respondents, the statistical relationships were so solid it was possible to identify fully 15 dimensions of personality. In the third pilot study, commercial intelligence test software was administered through a pocket game machine, the Nintendo 3DS, demonstrating how modern popular technologies can be adapted to radical purposes such as personality capture.

Other chapters explored the utility of recommender systems like the Netflix one, personalized expert systems, recording and reproducing personal autobiographic memories, natural language processing analysis of written text, and the use of online virtual worlds for both capture and emulation of personalities. All these studies were supported by references to the scientific literature, indicating how researchers working at the cutting edge in their fields are implicitly developing identity expansion technology, whether or not that is their goal. The book publisher is the same one that produced the most recent Converging Technologies report, and both can be obtained either electronically or on paper. An important indicator of the changing basis of human culture is the fact that 10 personality capture computer programs, used for several of the pilot studies in *Personality Capture and Emulation*, can be downloaded from the publisher's website for use by readers who wish to capture their own personalities. Progress in this field will require the creative energies of many scientists, engineers, and interested people in the general public, but they will experience payoffs every step of the way, as these new technological methods expand the scope of human action, awareness, and intellectual adventure.

4. Identity Transcendence

By any plausible demographic reckoning, most people who have ever lived are dead. Yet all their atoms still exist. Western religions tend to say that their souls migrated to an afterlife, while some eastern religions imagine they are periodically reincarnated, sans memories, in this world. But both assume that each person has a transcendent essence, call it the *immortal soul*, that can inhabit a physical body but is not limited to that form of existence. This is not merely a theological claim, that humans are like minor gods, but a way of conceptualizing the fact that our minds do not feel as if they were identical to our bodies. I suggest that this is a valid perception, but framed in primitive terms that may be considered obsolete in this post-modern age.

In a very real sense, a person is a dynamic pattern of information. Objectively, we are a particular configuration of atoms, not the atoms themselves. Subjectively, we are the data transmitted through our brain and wider nervous system, not the neurons themselves. Whatever value completely different conceptualizations may have, this one is in tune with our present era, in which every prosperous person owns a computer containing vast amount of

personally-relevant information, and often voyages out across Internet in one informatic form or another. If this conceptualization is factually correct, then the doctrine of immortal souls is not so much incorrect as it is primitive. An afterlife may exist, but in Cyberspace, not Heaven.

A different view emerged in the nineteenth century and is still developing today, based on evolution by natural selection from random biological variation. Existence consists of physical atoms distributed unevenly across space, and the atoms are sufficiently complex that some of them assembled through pure chance to form the basis of life. Most of the cosmos is dead, but on a very few rare planets like the Earth, primitive organisms evolved in the sea and eventually were complex enough to move onto the land. Within animal lineages, nervous systems evolved not with any intentional purpose, but as a result of the fact that they enhanced the survival and reproduction chances for the individuals who possessed them. Evolution is slow, and large populations have an evolutionary advantage because of their genetic diversity, so no one individual matters. The proper unit of analysis is the gene pool, a point made by Richard Dawkins in his popular book, *The Selfish Gene* (Dawkins, 1976; Strong & Bainbridge, 2003). Once very complex nervous systems had evolved, animals such as humans become aware of their own mortality, but the appropriate solution to the problem of death is reproduction, so that genes can live on in successive generations.

Yet another view has probably existed throughout history but was clearly articulated over the past couple of centuries. Humans are social beings. Without language we cannot think about issues of life and death, but languages were collectively constructed over hundreds of thousands of years in a social process. Thus, humans do not really exist as individuals but as societies. The best way for an individual to deal with death is to contribute so much to family, society and culture, that the deceased person will live on through the surviving people who benefitted from that person's actions. This may be what sociologist Émile Durkheim (1915) meant when he said that God is really a metaphor for society. It is certainly consistent with secularism and can supplement almost any religious doctrine.

We could call the evolutionary idea the *reproduction* solution to the problem of death, and the social one the *incorporation* solution, as a person surviving through good deeds becomes incorporated into the living community. Conceptualizing humans as dynamic patterns of information defines the *information* solution. None of these three requires supernatural

beliefs, and it would be philosophically consistent for a person who lacks faith to invest equally in the reproduction, incorporation and information solutions. That is to say, *convergence* of the three partial solutions is most practically effective and psychologically satisfying. At the present time, however, the information solution seems to offer the greatest potential for progress, so it is emphasized here.

Among the most controversial ideas is the intentional engineering of a new religious tradition promising immortality by means of interstellar delivery of digitized human personalities, perhaps using robots to colonize distant planets where traditional human bodies could be synthesized from DNA, or conceivably transmitting the data to extraterrestrial civilizations so that they can invite humans to dwell in the computers on their worlds (Bainbridge, 2011). Setting aside the myriad of technical challenges, the reason this idea is controversial is that religion is held sacred by believers, and thus universities do not teach religious engineering, when they logically could, applying the results of research in the psychology and sociology of religion. However, August Comte (1883), the founder of sociology, considered sociology itself to be a modern form of religion. Robert Geraci (2014) has studied the way Transhumanists have used the virtual world Second Life to achieve electronic transcendence, becoming a form of religion. At least one innovative religious "cult" that used electronic devices in its ministry has explicitly called its work "religious engineering" (Bainbridge 1978, p. 48). Giulio Prisco, a physicist and computer scientist, contributed a chapter on transcendent engineering to *The Transhumanist Reader*. He writes:

> I am persuaded that the ultimate realization of the dream of achieving an indefinite lifespan, with vastly enhanced cognitive abilities, lies in leaving biology behind and moving to a new, postbiological, cybernetic phase of our evolution. Mind uploading, the transfer of a human mind, memories, personality and "self" (whatever "self" is) to new high-performance substrates is the ultimate technology for immortality. Therefore I have always been interested in mind uploading and I consider it as the "Holy Grail" of transhumanism: let our minds break free of our biological brains and bodies, and we will be free to roam the universe and grow beyond limits as "software angels." (Prisco, 2013, p. 235)

The radical transformation of human beings suggested by Prisco is not merely a wild attempt to transcend the current human condition, but a plausible method of overcoming the practical barriers to interstellar

spaceflight. We have no difficulty sending space probes throughout the solar system, even beyond, and many of them have survived for years in environments where biological humans would die in an instant. Evolution has produced our intelligence within the confines of the terrestrial biosphere, and now that natural intelligence can create very different forms of artificial intelligence, suitable for a range of alien environments. If interstellar travel goes at the speed of Voyager II, it will take 50,000 years to reach another solar system, which would require unusually long-duration technologies, but would not require impossibilities.

The unification of the reproduction, incorporation, and information responses to death naturally follows the principles detailed in all the Converging Technologies reports, most specifically the convergence of biology, social science, and information science, assisted by cognitive science and nanoscience. The latest report however transcended even convergence itself, through a theory of the *convergence-divergence cycle.* As the executive summary of the new report explains (Roco et al., 2014),

Convergence is actually part of a dynamic and cyclical convergence–divergence process that originates organically from brain functions and other domains of the global human activity system. This process can provide a structure and specific improvement methods for the creative-innovation-production chain. The convergence phase consists of analysis, making creative connections among disparate ideas, and integration. The divergence phase consists of taking these new convergences and applying them to conceptual formation of new systems; application of innovation to new areas; new discoveries based on these processes; and multidimensional new outcomes in competencies, technologies, and products.

In biological evolution, the convergence-divergence cycle is manifested in speciation. In a large and diverse gene pool, many genes come together in a general configuration that is inherently new or well-adapted to new environmental conditions, converging to produce a new species, manifestly different from the old species and soon becoming reproductively separate from it. Then, if the new species is especially successful, it will undergo adaptive radiation, a divergence into many offspring species. In modern information systems, a diversity of people input data, and the data converge into a collective database. As people use the system, the data are personalized during the output phase, representing a form of divergence. With these two examples in mind, one can wonder how the convergence-divergence cycle might function

on a grand scale, if humanity is able to spread its diverse yet interacting cultures out across immensity to the stars.

Conclusion

On balance, Transhumanists have expressed more radical views than most participants in the Converging Technologies movement. However, both movements are optimistic that we are not fast approaching the limits of scientific discovery and technological invention (cf. Horgan, 1996; Barrow, 1998). Especially at the Converging Technologies conferences, one heard the view that innovation will stall quickly unless we energize widespread enthusiasm for research and development, even if that progress faces steeply increasing investment costs. Only time will tell whether critics of the two parallel movements were right, that progress is coming inexorably to an end. Yet one of the best arguments for optimism is that pessimism is a self-fulfilling prophecy.

The most optimistic scenario would assume that both the Argument from Design and the Anthropic Principle are true, but for different periods of human history. The Anthropic Principle explains the past prior to the Omicron Point, by natural selection from random events. The Argument from Design explains the future after the Omicron Point, but in a novel way. God did not design the universe in the past, because he never existed. But having imagined the concept of god, we can play that role, creating a universe that has purpose. To accomplish that goal, we would first need to transform ourselves into gods.

REFERENCES

Bainbridge, W.S. (1978). *Satan's Power: A Deviant Psychotherapy Cult.* Berkeley: University of California Press.

Bainbridge, W. S. (1991). *Goals in Space: American Values and the Future of Technology.* Albany, New York: State University of New York Press.

Bainbridge, W.S. (1997). The Omicron Point: Sociological application of the anthropic theory. In Eve, R.A., Horsfall, S. & Lee, M.E. (eds.), *Chaos and Complexity in Sociology: Myths, Models and Theory.* Thousand Oaks, California: Sage.

Bainbridge, W.S. (2004a). Progress toward cyberimmortality. In Klein, B.J. & Sethe, S. (eds.), *The Scientific Conquest of Death*. Buenos Aires: Libros En Red, 107-122.

Bainbridge, W.S. (2004b). The evolution of semantic systems. In Roco, M.C. & Montemagno, C.D. (eds.), *The Coevolution of Human Potential and Converging Technologies*. New York: New York Academy of Sciences, 150-177.

Bainbridge, W.S. (2006a). Cognitive technologies. In Bainbridge, W.S. & Roco, M.C. (eds.), *Managing Nano-Bio-Info-Cogno Innovations: Converging Technologies in Society*. Berlin: Springer, 203-226.

Bainbridge, W.S. (2006b). Transformative concepts in scientific convergence. In Bainbridge, W.S. & Roco, M.C. (eds.), *Progress in Convergence: Technologies for Human Wellbeing*. New York: New York Academy of Sciences, 24-45.

Bainbridge, W. S. (2007). *Nanoconvergence*. New Jersey: Prentice-Hall.

Bainbridge, W.S. (2009). Personality enhancement and transfer. In Healey, P. & Rayner, S. (eds.), *Unnatural Selection: The Challenges of Engineering Tomorrow's People*. London: Earthscan, 32-39.

Bainbridge, W.S. (2010). Burglarizing Nietzsche's tomb. *Journal of Evolution and Technology*, *21*(1), 37-54.

Bainbridge, W.S. (2011). Direct contact with extraterrestrials via computer emulation. In Vakoch, D.A & Harrison, A.A. (eds.), *Civilizations Beyond Earth: Extraterrestrial Life and Society*. New York: Berghahn, 191-202.

Bainbridge, W. S. (2013a). *eGods*. New York: Oxford University Press.

Bainbridge, W. S. (2013b). Transavatars. In More, M. & Vita-More, N. (eds.). *The Transhumanist Reader*. Chichester, West Sussex, UK: Wiley-Blackwell, 91-108.

Bainbridge, W. S. (2014). *Personality Capture and Emulation*. London: Springer.

Bainbridge, W.S. & Roco, M.C. (eds.). (2006a). *Managing Nano-Bio-Info-Cogno Innovations: Converging Technologies in Society*. Berlin: Springer.

Bainbridge, W.S., & Roco, M.C. (Eds.). (2006b). *Progress in Convergence: Technologies for Human Wellbeing*. New York: New York Academy of Sciences.

Barrow, J.D. (1998). *Impossibility: The Limits of Science and the Science of Limits*. Oxford, England: Oxford University Press.

Barrow, J.D. & Tipler, F.J. (1986). *The Anthropic Cosmological Principle*. New York: Oxford University Press.

Bergson, H. (1911). *Creative Evolution*. New York: H. Holt.

Bertocci, P.A. (1945). Teleological argument for God. In Ferm, V. (ed.), *An Encyclopedia of Religion*. New York: Philosophical Library, 763.

Carr, B.J., & Rees, M.J. (1979). The anthropic principle and the structure of the physical world. *Nature*, *278*, 605-612.

Comte, A. (1883). *The Catechism of Positive Religion*. London: Trübner.

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

Durkheim, É. (1915). *The Elementary Forms of the Religious Life*. New York, Free Press (1965).

FM-2030. (1989). *Are You a Transhuman?* New York: Warner.

Fukuyama, F. (2002). *Our Posthuman Future: Consequences of the Biotechnology Revolution*. New York: Farrar, Straus, and Giroux.

Fukuyama, F. (2004). The world's most dangerous ideas: Transhumanism. *Foreign Policy*, September/October, 144, 42-43.

Gale, G. (1981). The anthropic principle. *Scientific American*, *245*(6), 154-171.

Geraci, R.M. (2014). *Virtually Sacred*. New York: Oxford University Press.

Henderson, L.J. (1913). *The Fitness of the Environment*. New York: Macmillan.

Henderson, L.J. (1917). *The Order of Nature*. Cambridge: Harvard University Press.

Horgan, J. (1996). *The End of Science: Facing the Limits of Knowledge in the Twilight of the Scientific Age*. Reading, Massachusetts: Addison-Wesley.

Hughes, J. (2006). Human enhancement and the emergent technopolitics of the 21st century. In Bainbridge, W.S. & Roco, M.C. (eds.), *Managing Nano-Bio-Info-Cogno Innovations: Converging Technologies in Society*. Berlin: Springer, 285-308.

Kurzweil, R. (1999). *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*. New York: Viking.

Kurzweil, R. (2004). Human body version 2.0. In Klein, B.J. & Sethe, S. (eds.), *The Scientific Conquest of Death*. Buenos Aires: Libros En Red, 93-106.

Leslie, J. (1982). Anthropic principle, world ensemble, design. *American Philosophical Quarterly*, *19*, 141-151.

Miah, A. (2006). Rethinking enhancement in sport. In Bainbridge, W.S. & Roco, M.C. (eds.), *Progress in Convergence: Technologies for Human Wellbeing*. New York: New York Academy of Sciences, 301-320.

Minsky, M.L. (2004). Will robots inherit the Earth? In Klein, B.J. & Sethe, S. (eds.), *The Scientific Conquest of Death*. Buenos Aires: Libros En Red, 123-134.

Minsky, M.L. (2006). *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. New York: Simon and Schuster.

More, M. (2013). The philosophy of Transhumanism. In More, M. & Vita-More, N. (eds.). *The Transhumanist Reader*. Chichester, West Sussex, UK: Wiley-Blackwell, 3-17.

More, M., & and Vita-More, N. (2013). *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*. Chichester, West Sussex, UK: Wiley-Blackwell.

Nietzsche, F. (1872). *Die Geburt der Tragödie*. Munich: Goldmann.

Nietzsche, F. (1885). *Also Sprach Zarathustra*. Stuttgart: Kroner.

Prisco. G. (2013). Transcendent engineering. In More, M. & Vita-More, N. (eds.). *The Transhumanist Reader*. Chichester, West Sussex, UK: Wiley-Blackwell, 234-240.

Roco, M.C., & Montemagno, C.D. (eds.). (2004). *The Coevolution of Human Potential and Converging Technologies*. New York: New York Academy of Sciences.

Roco, M.C. & Bainbridge, W.S. (2001). *Societal Implications of Nanoscience and Nanotechnology*. Dordrecht, Netherlands: Kluwer.

Roco, M.C. & Bainbridge, W.S. (2003). *Converging Technologies for Improving Human Performance*. Dordrecht, Netherlands: Kluwer.

Roco, M.C. & Bainbridge, W.S. (eds.). (2006b). *Nanotechnology: Societal Implications--Maximizing Benefit for Humanity*. Berlin: Springer.

Roco, M.C., & Bainbridge, W.S. (eds.). (2006a). *Nanotechnology: Societal Implications--Individual Perspectives*. Berlin: Springer.

Roco, M.C., Bainbridge, W.S., Tonn, B. & Whitesides, G. (eds.). (2014). *Convergence of Knowledge, Technology, and Society: Beyond Convergence of Nano-Bio-Info-Cognitive Technologies*. Berlin: Springer.

Sandberg, A., & Bostrom, N. (2006). Converging cognitive enhancements. In Bainbridge, W.S. and Roco, M.C. (eds.), *Progress in Convergence: Technologies for Human Wellbeing*. New York: New York Academy of Sciences, 201-227.

Sententia, W. (2004). Neuroethical considerations: Cognitive liberty and converging technologies for improving human cognition. In Roco, M.C. and Montemagno, C.D. (eds.), *The Coevolution of Human Potential and Converging Technologies*. New York: New York Academy of Sciences, 221-228.

Sententia, W. (2006). Cognitive enhancement and the neuroethics of memory drugs. In Bainbridge, W.S. & Roco, M.C. (eds.), *Managing Nano-Bio-Info-Cogno Innovations: Converging Technologies in Society*. Berlin: Springer, 153-172.

Sorgner, S.L. (2009). Nietzsche, the Overhuman and Transhumanism. *Journal of Evolution and Technology, 20*(1), 29-42.

Stapledon, O. (1935). *Odd John*. London: Methuen.

Strong, G., & Bainbridge, W.S. (2003). Memetics: A potential new science. In Roco, M.C. & Bainbridge, W.S. (eds.), *Converging Technologies for Improving Human Performance*. Dordrecht, Netherlands: Kluwer, 318-326.

Sturgeon, T. (1953). *More than Human*. New York: Farrar, Straus and Young.

Teilhard de Chardin, P. (1964). *The Future of Man*. New York: Harper.

Van Vogt, A.E. (1946). *Slan*. Sauk City, Wisconsin: Arkham House.

White, H.C., Boorman, S.A. & Breiger, R.L. (1976). Social structure from multiple networks: Blockmodels of roles and positions. *American Journal of Sociology*, *81*(4), 730-780.

Wowk, B. (2013). Medical time travel. In M. More & N Vita-More (Eds.). *The Transhumanist Reader*. Chichester, West Sussex, UK: Wiley-Blackwell, 220-226.

Cognitive Enhancement and Personal Identity

Roberto Mordacci [†]
mordacci.roberto@unisr.it

ABSTRACT

Enhancing cognition is a complex activity, for the sake of which humanity has developed a rich array of techniques and skills. We can distinguish between three categories: a) cognitive supports and education; b) neural cognitive enhancers: drugs and other ways to improve the functionality of cognitive neural networks; c) technological cognitive enhancers: implants, extended minds and technological supports variously integrated in the neural cognitive networks. Applying a version of the Parity Principle, I argue that there is no morally relevant difference in the three categories. What we want to preserve while using these techniques is not the biological status quo of the mind of persons, but rather personal identities. In this perspective, there can be no general objection to cognitive enhancement. Every technique, even very traditional ones, have their drawbacks, especially when they threaten to reduce the autonomy of agents in moulding their own personal identity.

Introduction

In this paper, I would like to propose a rather general argument in favour of cognitive enhancement. Yet, at the same time, I suggest that we should consider cognitive enhancement as a part of a more general issue, i.e. personal identity (or, as I would rather say, individual personality) as the result of a set of practices, actions and choices by which we define who we are through our practices. These practices and their connection with our personal identity offer a criterion for evaluating particular enhancing techniques.

[†] Università Vita-Salute San Raffaele, Milano, Italy.

What is cognitive enhancement? It is rather difficult to give a straightforward definition, since cognition is an activity which involves a vast array of practices, means and social systems. In the history of the debate, quite narrow definitions of cognitive enhancement were initially given:

> Interestingly, the term 'cognitive enhancement' was originally used to describe the treatment of disease-associated cognitive impairment, such as in dementia and schizophrenia, and involved using various strategies to boost cognitive functions. The meaning of the term was subsequently broadened to encompass the use of interventions for mild cognitive impairment (MCI), currently defined as cognitive deficits that do not overtly impair function. Nowadays 'cognitive enhancement' is often applied exclusively to interventions in normal ageing processes and in 'healthy' people for non-medical purposes. (Ferrari, Coenen, Grunwald 2012, p. 220)

It is now common usage to understand cognitive enhancement in a wider perspective, and particularly not only in relation to health or illness. The process of cognition is extremely complex, and it must be stimulated if it is to be efficient according to our needs in the context we find ourselves in. Conceptually, we have good reasons to consider as "cognitive enhancement" *any activity that fosters our basic cognitive abilities.*

Cognition is not just a simple act performed by an individual: this would be a rather reductive view of the matter. Cognition is a process whereby information is acquired, selected, memorized and put to use in the pertinent setting. A complex interplay of abilities and means are involved in this process and it can be said that this is at the same time an individual and a social activity: a reasonable validation of one's cognitive acts often requires some kind of validation from other agents, just like when, in front of a strange phenomenon, I wonder: "Should I believe what I see?". In a case like this, I usually look for other agents to confirm or disconfirm my impression, and although it can happen that we are all deluded, still we consider visual testimony as a sufficient proof of real events in our everyday life.

Gathering information has always been a social activity, where direct personal experience is always exposed to limits and where I do need to trust, at least minimally, the information given to me by other people (that's why the virtues of Accuracy and Sincerity are so important, as Bernard Williams has warned us; see Williams 2002). Thus, even if single acts of cognition are the works of individuals, cognition in general is a socially embedded activity, especially as we move away from a very basic condition of elementary knowledge.

To put it shortly, *cognition may be conceived of as an individual activity in a social environment.* Exchanging information is one of the main ways to get to know something, which is the first step of cognition. In comparison with other species, human beings have developed the ability of non-genetic learning, which has detached information from the body and made it quickly available to other individuals through communication and education. This is the main way by which human knowledge has improved since the raising of the Homo Sapiens (Williams, 2002).

So what does it mean to enhance cognition? In the first place, it should mean to enhance the ability of individuals in a social context to acquire, select, memorize and make use of information. Therefore, cognitive enhancement can be understood as the increment of our ability as individuals to acquire, select, store and use knowledge.

Basically, this is an activity in which we as a species are involved since our appearance on earth. At the individual level, our ability of acquiring and storing information for present or future use is enhanced by a number of possible means, from memorization techniques to biotechnological implants: before asking the moral question about cognitive enhancement, it can be useful to distinguish some categories of means that we employ for this goal.

1. Three Categories In Cognitive Enhancement

For the sake of the argument that I want to develop here, we might recognize three categories of cognitive enhancers.

Cognitive Supports and Education: There are traditional means for developing (enhancing) cognitive abilities: books, schools, higher education, courses, meetings, libraries, archives, and today databases, Internet encyclopedias, google and other search engines, and even forums and social networks. None of these is an enhancement *per se*: if you do not use them appropriately, your cognitive abilities remain just the same as before. A book does not enhance cognition if you do not read it; and sometimes even when you do, especially if it is a book of philosophy.

These means need an effort by the individual, usually in connection with a socially established training practice. Therefore, they are more appropriately defined as cognitive supports utilized within activities related to education and training. Furthermore, these means are in quite a clear sense *external* to the individual: they enhance cognition as the result of the training of a

natural endowment. Yet, we should be aware that these means do enhance cognition only if methods and information are in an important sense *internalized*, i.e. made part of the individual. We usually call this rather artificial endowment "culture" or education (in the sense of some special training) or "being well read".

These means of cognition are widely supported and valued worldwide. It is deemed noble to improve one's knowledge and it is considered debatable not to cultivate one's cognitive talents. Apart from the education of kids, who need to develop basic cognitive abilities and for whom it has been a long and still unfinished fight to ensure a fair access to schooling, even the most traditional means of cognitive enhancement are unevenly distributed and used to be very elitist. Criticism against restricted access to higher education were raised only in a relatively recent time and had a burst out in the 1970s, while they seem to fade more recently. Nowadays, it seems that access to universities has been made less easy, also in economic terms. Apart from this, cognitive competition has been traditionally valued as an enhancer for students. In general, we tend nowadays to speak about lifelong education, meaning that we still think that to improve one's knowledge is a lifelong duty and an opportunity. As for issues of justice, it is not cognitive equality which is looked for by general educational programs, but rather fairness of access to the means of cognitive improvement offered by higher education.

There have been critiques to highly specialized training, even in science: an excessively restricted scope of interest in an area of knowledge is said to make persons rather blind to general issues and, in terms of personal identity, rather prone to give up other moral and psychological features in their character for the sake of science.

In this sense, even traditional supports to cognitive abilities have been subject to criticism. To sum up, these means of enhancing cognition show pros and cons.

Pros: schooling, learning and specialized training are easy to share, they can enter into public programs and be fairly distributed among the population. They are based on methods and contents and enter into the constitution of the individual as tools for the construction of one's personality. They are based on human relationships such as that of teacher and student, or expert and trainee.

Cons: these means require extensive and expensive public programs in order to be effective on a large scale. The rate of failure of the process depends on many factors: the effort of the students, the ability of teachers, the validity of

the methods. Books and science are not always accurate as they should. The individual is often put under a very heavy stress in order to enter the competitive world of knowledge. A personality entirely devoted to scientific knowledge is not necessarily a flourishing one.

Neural Cognitive Enhancers: A second category of enhancers can be called "Neural Cognitive Enhancers": I enclose here various ways to improve the functionality of cognitive neural networks in their biological status. They are mainly drugs used therapeutically to treat syndromes which influence the ability to pay attention and stay alert, such as ADHD and narcolepsy. The drugs are mainly methylphenidate and modafinil; in some cases, beta blockers are used as well. In these cases, these drugs have proven effective in ameliorating the cognitive functions of patients. There is some evidence (though its meaning is disputed) that these drugs are used without medical prescription by university students in order to enhance their cognitive abilities (Rabiner et al., 2009). The percentage of students taking these drugs is not very high and it seems it is fading in recent years. Furthermore, there is no convincing evidence that they work as enhancers of the performance of students, who seem to be motivated by the goal of obtaining higher scores. Yet, it is disputable whether any kind of real enhancement is going on here or not: many of the students using these drugs were reported as showing signs of an undiagnosed attention deficit disorder or other cognitive problems. Using these drugs they do perform better than they used to, but they do not perform better than the most brilliant students, who do not seem to make use of these drugs. So it seems that the motivation is rather a struggle for equality rather than enhancement. So, it has been noted that if such motivations are indicative of self-treatment we could expect the baseline of academic success of students taking cognitive enhancers to be below average due to their undiagnosed cognitive deficits by comparison to the average student body. Indeed, this expectation, although tentative, has been suggested by Rabiner et al. (2009) who found that students engaging in cognitive enhancement did indeed have lower than average academic scores and were thus struggling academically in comparison to the main student body. (Outram, 2012, p. 177)

Therefore, it may be suggested either that we revise the expectation that such non-medical use of stimulants for academic purposes is necessarily cognitive enhancement on the basis of being undertaken by truly healthy individuals, or we reflect upon the difficulty of separating self-treatment from enhancement. (Outram 2012, p. 177)

It can therefore be argued that «it should not be assumed that all forms of nonmedical stimulant use are necessarily forms of cognitive enhancement» (Outram 2012, p. 180). If we accept a more restricted definition of cognitive enhancement (in this category) as «the use of drugs and other interventions to modify brain processes with the aim of enhancing memory, mood and attention in people who are not impaired by illness or disorder» (Hall, W. 2004. Feeling 'better than well'. EMBO Reports 5 (12): 1105–1109) we might question that so far we have a convincing evidence that healthy students use cognitive enhancement drugs and that, even if they do, that it is effective in enhancing their cognitive abilities. (Perhaps the really smart ones do not need enhancers or they do not want to use them).

Of course, these drugs do not substitute for personal efforts and hard work in gathering, selecting, storing and using information: traditional ways of gaining knowledge are simply made somewhat easier. In short terms, a more focused attention naturally ends up in better memorized data. The real challenge is to integrate these information in a *body of knowledge* available to the individual in the circumstances where it is needed. There is no evidence that this happens, since even the performance of a student is measured rather on single tasks (examinations) than on a lifelong competence.

An objection that is quickly raised against the use of these drugs, especially in students (but in scholars as well), might be more of a psychological rather than of a moral kind: if the environment is highly competitive, I will need to be always at my more-than-best in order to survive and win competitions. Now, if I believe that my performance depends on stimulants (and maybe not only cognitive ones), it is very likely that I will develop a dependence on them. Or, at least, I will be inclined to believe that I am not adequate to my environment unless I take these enhancers. There is little evidence on side effects of these drugs, but this kind of dependence is not of the organic kind.

This objection must be clearly distinguished from the so-called objection of "inauthenticity", i.e. that the drugs make me somewhat "different from what I am" – as it is sometimes said of antidepressants (e.g. Prozac) used off-label as mood stimulants. Now, to be clear on this point: there is no "original self" under threat here, since in the (Kantian) perspective I am assuming the self is an ongoing construction, not a given and not anything which precedes action. The objection is rather that building one's personal identity relying heavily on stimulants or enhancers may offer good results in single performances but has the drawback of projecting my self-image as that of a less-than-adequate person

when enhancers are not available. (It is like when super-heroes lose their powers: they immediately fall prey of a dire identity crisis).

So, we can sum up pros and cons like follows:

Pros: there is no convincing evidence, but students using these drugs do report increased attention and memory. It is to be seen whether this improves creativity and the ability to trace connections between pieces of information, so that what is learned is turned into a reasonably stable body of knowledge. In this case, it is certainly a form of cognitive enhancement, which per se cannot be but good. The data are missing, though.

Cons: The risk of psychological dependence and underestimation of oneself in the absence of enhancers should not be dismissed too easily. After all, we look for improved cognition in order to be better persons (see below). The effects of these enhancers seem to be short-termed, and still an effort on the part of the individual is required. Concentration is enhanced, but it is not clear that knowledge is in the end incremented. We do not know what the real effects of long-term use are, in terms of physical effects, psychological dependence and efficacy in extending the knowledge available to the individual.

(Bio-)Technological Cognitive Enhancers: Drugs seem to be able to alter the functionality of neurons without altering the biological status of the brain. Some other procedures are more invasive. I am thinking here of technological supports which can be variously integrated in the autonomous functioning of neural cognitive networks. Among these we can further distinguish two subcategories: a) devices which simply directly stimulate the brain (deep brain stimulation, transmagnetic stimulation, genetically modified neural cells); b) one can imagine of neural implants, microchips, extended minds and any kind of hardware added to the brain as an extension of memory or a further PCU. There is a lot of speculation when we enter this third category of enhancers.

With respect to deep brain stimulation and transmagnetic stimulation, we have little or no empirical evidence of anything that could be called a widespread use of these technologies by members of the public wanting to enhance their abilities. Indeed, the British Medical Association report on cognitive enhancement has highlighted, such technology is largely experimental and it is "highly questionable whether healthy people would want, or should be encouraged to want, to have invasive brain surgery, with all its attendant risks, in order to enhance their cognitive ability". In a similar vein, concerning transmagnetic stimulation, the report declares that "[a]lthough

research has identified some small, short-lived, task specific improvements in a laboratory setting, this is very different from the significant, long-term, useful improvements that would be required to justify its use in real-life settings and on a population basis". (quoted in Outram 2012, pp. 174-175)

Biomedical cognitive techniques include the administration of drugs, implants of genetically engineered or stem-cell-grown neural tissue, transcranial magnetic stimulation, computer/brain interfacing (already used to simulate vision and enable movement in people with severe neurological damage), and (perhaps someday) the application of genetic engineering and/or synthetic biology methods to human embryos or gametes. (Buchanan, 2011, p. 146).

The border between this kind of enhancers and other devices that we normally use today (e.g. smartphones) and will probably use even more in the future is difficult to trace clearly. Basically, the difference is in that these biotech enhancers are integrated in the "normal functioning" of the brain, while present devices still seem to be detached from our body. I do not want to enter here the issue of extended minds and the self, but it is clear that if we consider cognitive enhancement in a sufficiently wide perspective, we have to consider the following issue: to what extent do we consider ourselves as defined by our degree of knowledge as individuals?

In terms of common sense psychology, our cognitive endowment is as large as the information available to us in a reasonably short time (on the analogy of working with our existing memory and associative abilities). But it is becoming more and more common for us to consider the memories of our external devices as (easily reachable) parts of our memory.

Taking the famous example of Otto and Inga, most of us are more or less in the middle: we do remember a lot of things about ourselves and our town, but who is not using googlemaps when she does not remember exactly how to reach the Metropolitan Museum or when she is looking for the nearest affordable restaurant ("yes, I've visited one not far from here a couple of months ago, but where it was precisely and how was it called?")?

Biotech implants, although only imagined so far, are thought of as internal to the individual, integrated in the normal functioning of the brain. This is why they are not subject to the objection of dependence: if we imagine them as technological devices which, one day, might take from the body itself the energy they need to work, there is no dependence here, apart from failures which are analogous to illnesses. Expanding our memory is certainly a

cognitive enhancement, and if this can be done without impairing other brain functions there seems to be no reason to object to it.

2. A Parity Principle Between Traditional And Non-Traditional Means

The strong version of the Ethical Parity Principle (sEPP) introduced by Neil Levy says that «alterations of external props used for thinking are (ceteris paribus) ethically on a par with alterations of the brain» (Levy, 2007). The formulation is a bit strange, since we are likely to use the Principle in order to judge which alterations of the brain are acceptable on the basis of the acceptability of the alterations of the external props and not vice-versa. The weak version of the Principle (wEPP) makes this almost explicit when it says that «Alterations of external props are (ceteris paribus) ethically on a par with alterations of the brain, to the precise extent to which our reasons for finding alterations of the brain problematic are transferable to alterations of the environment in which it is embedded». The emphasis on reasons points to the fact that what we think ought to be protected while altering our cognitive abilities is not so much the fact that alterations take place inside or outside the brain. The reasons for refusing an alteration of the brain are connected to the fact that we value cognition as a part of a more general and hierarchically superior value, i.e. the value of the person. If we could obtain outstanding results in cognitive processes at the price of devastating other functions of the brain which are essential to the normal functioning of the person, we would probably object to it. And the reason to object would bear some similarity to the Kantian principle that persons are always to be considered as ends and not as mere means.

We do think that cognition is an important part of our personality, and many of us do ground our self-esteem on the basis of the knowledge we can display exactly when needed. Nonetheless, we still tend to refuse a total identification of ourselves with our knowledge: we are not willing to pay any price for it.

Now, drawing on this analogy and on the reported differences, I would suggest, as a principle for evaluating cognitive enhancement techniques, the following Cognitive (Weak) Parity Principle:

Cognitive (Weak) Parity Principle: alterations of the cognitive processes which take place inside the brain are ethically on a par with alterations of cognitive processes taking place outside the brain, in so far as the whole person is not

damaged in her other functions and in the construction of her personality (personal identity).

Since our cognition, as I said before, is a widely social activity, we have grounds for objections against those activities which, destroying parts of our cognitive endowment and our educational system outside the brain, make it difficult for individuals to improve their cognitive abilities through traditional means. Insofar as these process are internalized and made easier through pharmacological and biotechnological means, there seems to be no objection in principle against this, provided that:

1)   The (enhanced) cognitive processes are capable of being integrated in the body of knowledge that the individual can use when needed and in the framing of her personal identity.

2)   The processes themselves do not pose threats, in the enhanced status, to the ability of the person of developing an autonomous sense of herself and an adequate self-image. This might happen, for example, if the enhanced cognitive processes imply some reduction or distortion of, e.g., the normal emotional or relational abilities of persons

Conclusion

As a conclusion, we may consider a general point: the discussion on cognitive enhancement should probably be set against the background of the meaning of knowledge and cognitive abilities for the life of individuals and for the value we as a society assign to those abilities. For example, we are not so keen on cognition that we would plainly accept the use of pharmacological cognitive enhancers if they have important side effects and are not very effective.

Those who argue in favour of liberalising pharmaceutical cognitive enhancers, for example, would do better to ask whether we as a society are ready to accept the consumption of pharmaceutical substances whose effects have not been fully tested (and, if so, why we are prepared to accept lower safety standards for enhancers in healthy subjects than in established standards for therapeutic uses on patients), rather than whether banning these enhancers is compatible with our respect for autonomy (cf. [19]). (Ferrari, Coenen, Grunwald 2012, p. 227).

Yet, autonomy is indeed a value and, provided that the standards of efficacy and safety are respected, we generally give value to autonomous decisions. The point is that it is not only autonomous decision that we value: we look for a

shape of ourselves that reflects our complexity and keeps it in a somewhat harmonic unity. The fundamental criterion for evaluating cognitive enhancement is its relation to our striving for personal flourishing, which is something different from just acting autonomously or being free to use whatever means in order to do what we want. We look for those goods, cognition among others, which make our identity a construction to which we can give a meaning for us and for others. The real obstacle to the whole debate is the idea that the issue is whether cognitive enhancements threat some hypothetical "natural" or "original" self, hidden somewhere in the status quo of our abilities. As Allen Buchanan has written,

> Given a plausible understanding of molecular-developmental and evolutionary biology, the cognitive potential that human beings typically have is not unalterable and not likely to be optimal. Pursuing the goals of education may require changing what we have hitherto regarded as the individual's 'natural' potential, even in the case of normal individuals, and this, in turn, may require recourse to Biomedical Cognitive Enhancement. (Buchanan, 2011, p. 147)

What we really care for is the possibility of innerving our abilities with a sense of personal presence, the adherence of our dynamic capacity for realizing good things to the complex of our forces, energies, innate abilities (whatever they are) and acquired capacities. And we want that our attempt is in principle understandable by any other and, hopefully, even approved of, or even appreciated, praised.

So, enhanced cognitive abilities can of course be valued and appreciated. They are, when they are developed through traditional means. And if newer, effective and safe means offer the possibility of integrating those empowered abilities into our comprehensive self, as autonomous agents committed to the construction of a recognizable identity, then the moral point of view should not be hostile to them.

REFERENCES

British Medical Association (2007). Boosting your Brainpower: Ethical Aspects of Cognitive Enhancement. available from http://www.bma.org.uk/-images/Boosting_brainpower_tcm41-147266.pdf, [accessed 21 February 2011].

Buchanan, A. (2011). Cognitive enhancement and education. *Theory and Research in Education 9*(2), 145-162.

Ferrari, A., Coenen, C., Grunwald, A. (2012). Visions and Ethics in Current Discourse on Human Enhancement. *Nanoethics 6*, 215–229.

Hall, W. (2004). Feeling ′better than well′. *EMBO Reports 5(12),* 1105–1109.

Levy, N. (2007). *Neuroethics. Challenges for the 21st Century*, Cambridge, MA: Cambridge University Press

Outram, S.M. (2012). Ethical Considerations in the Framing of the Cognitive Enhancement Debate. *Neuroethics 5,* 173–184.

Rabiner, D.L., A.D. Anastopoulos, E. Jane Costello, R.H.Hoyle, S.E. McCabe, and H. Scott Swartzwelder (2009). Motives and perceived consequences of nonmedical ADHD medication use by college students. *Journal of Attention Disorders 13(3)*, 259–270.

Williams, B. (2002). *Truth and Truthfulness: An Essay in Genealogy.* Princeton: Princeton University Press.

Girl, Pixelated – Narrative Identity, Virtual Embodiment, and *Second Life* *

Anna Gotlib [†]
agotlib@brooklyn.cuny.edu

ABSTRACT

This paper focuses on the reasons for, and consequences of, expanding our notions of human embodiment to virtual worlds. Increasingly, it is within virtual environments that we seek to extend, and enhance, who we are. Yet, philosophical worries persist about what sorts of selves count as moral agents, and the extent to which self-enhancements affect personal identity and agency. This paper critiques and expands the discourse on embodiment and personal identity by locating it within the virtual environments of *Second Life*, challenging the prevailing limitations of what counts as identity-constituting embodiment. I argue that more inclusive notions of embodiment make possible a deeper understanding of its moral and epistemic force that constitutes and locates our identities in a universe of shared moral understandings. Thus, by including enhanced virtual embodiments alongside the non-virtual, not only do we expand our ideas of what it might mean to be embodied, but we also deepen our moral vocabularies of the self.

> ...I'd think, That ain't me, that ain't my face. It wasn't even me when I was trying to be that face. I wasn't even really me then; I was just being the way I looked, the way people wanted. It don't seem like I ever have been me.
> — Ken Kesey, *One Flew Over the Cuckoo's Nest*

Introduction

In *Descartes' Error: Emotion, Reason, and the Human Brain,* Antonio Damasio has famously argued against the separation of mind, body, rationality and emotions, and, as a corollary, for the primacy of embodiment in human experience, claiming that "[n]ature appears to have built the apparatus of rationality not just on top of the apparatus of biological regulation, but also *from* it and *with* it" (Damasio, 1994, p. 128). Others, including Merleau-Ponty, have agreed, noting that our embodied contacts with the world, and with other human beings, always precede the linguistic and reflective treatment of these encounters: We are often delayed in our accounts of what happened, always piecing together stories, reasons, and language about our experiences after the fact, while our embodied lives, whatever else they might be, are always firmly grounded in the now -- in the unrefined immediacy of the encounter between our bodily senses and the outside world. Gradually, views began to emerge that took seriously the proposition that we are subjects embedded in environments and enmeshed in contexts, and that embodiment was not something to be considered apart from our self-concept and perception -- or, indeed, in conflict with them -- but as an unavoidable, fundamental fact about who we are, and how we develop our reasons, desires, and selves.

Yet this gradual openness to the body as an intrinsic part of the self, and to embodiment as a moral, rather than just a physical, notion, also serves as a reminder that the body remains a contested site. Serious worries about what sorts of bodies "count" as "legitimate," how they are perceived, and how they are psychosocially located has occupied many of the feminist, queer, disability, race, and other theorists, who have argued for a broader, deeper, and more nuanced understanding of burdened embodiments. This paper, although quite sympathetic to this work, attempts to extend, and thus re-theorize, embodiment as a locus of self in a way that transforms and challenges its existing boundaries.

Specifically, I focus on the reasons for, and consequences of, expanding our notions of gendered embodiment to virtual worlds. My claim is made against the aforementioned background of serious worries about how contested bodies are perceived, and the effect this perception has on personal identity. Specifically, I argue that our ideas of a world in which we are embodied ought to include the virtual environments where increasingly large numbers of us spend our time. I suggest that virtual worlds, such as *Second Life,* challenge

the theoretical limitations of what counts as identity-constituting embodiments. The central questions that this paper asks are: (1) How does one sense one's pixelated body as one's own, and how does this experience contribute to one's sense of who one is as an embodied being? (2) Specifically, what do virtual embodiments offer to the identity-constituting narratives of those who might be in some ways "othered" in the non-virtual world? and

(3) What is the significance of these findings for our theorizing about the relationships between the body and one's identity? I claim that by allowing virtual embodiments to be considered alongside non-virtual ones as constitutive parts of one's identity, not only do we expand our notions of what it might mean to be an embodied being, but we also deepen our moral vocabularies of the self. This paper, then, is about taking new worlds, and our embodiment in them, seriously.

1. Narrative and The Embodied Self

Given that this paper is premised on the notion that embodiment is a fundamental fact of who we are, how we perceive, and how we are perceived, I begin with an examination of embodiment's narrative structure, and suggest that it is this narrative quality that connects it to identity formation. I take this initial scene-setting to be essential for three reasons: First, it provides an overview of what I take narrative embodiment to mean. Second, it demonstrates the connection between embodiment and personal identity. Finally, it provides a descriptive and evaluative language and a conceptual vocabulary that just might offer us a way to re-imagine embodiment.

What the concept of embodiment is -- or, more precisely, what they are -- is a complicated question. In "What Makes a Body?",Mark Johnson distinguishes five dimensions of human embodiment: (1) biological; (2) ecological; (3) phenomenological; (4) social; and (5) cultural. He argues that although they are not readily reducible to each other, all of these varieties of embodiment are interrelated and interdependent (Johnson, 2008, p. 164-166). I suggest that the way in which they come together in a self (or, indeed, selves) is through what might most accurately be understood as a narrative process. I begin with two fundamental questions: what is the embodied self, and how is that embodied self narratively constituted?

Although this paper problematizes what I suggest are new, and important, kinds of embodiment, the notion of human embodiment as a normative

concept that is central to concerns about identity and the self is at the center of a rich and broad discourse that incorporates ethical, metaphysical, epistemic, and sociopolitical dimensions (See Parfit, 1984; Ricouer, 1991; Merleau-Ponty, 1992; Lindemann, 1997; Meyers, 1997; Atkins, 2000). The claim is this: we live as, with, and through, our bodies. We think, make decisions, and act in ways that express our, and others', understandings of ourselves as embodied beings in a sociophysical universe. Indeed, Merleau-Ponty has argued that perception and consciousness (as well as other aspects of experiencing) are not abstract, theoretical constructs, but are reflections of a world understood and expressed through the body. The self is thus necessarily an embodied one in that our bodies are neither detached from experience, nor secondary to it. Simply, they are a part of the world in which we find ourselves (Merleau-Ponty, 1992).

This seems largely right. Indeed, it is through my body's sensory and motor capacities that first I encounter the world and understand myself as a particular being, as me, reflecting and reasoning about it only as a result of such encounters (Atkins, 2000). It is through my body that I first learn hot and cold, light and dark, and often, love and its absence. To the extent that I am able, I situate myself in the world as a particular kind of body, with some limitations that are both shareable with others, and limitations that are unique, private, or not readily visible.

Indeed, and despite Derek Parfit's protestations, it matters to me what happens to my body, that it continue to exist, and that it occupy certain positions in the world (say, that of someone who is not tortured) and not occupy others (say, that of someone forced into sex trade) (Parfit, 1984). Thus, as my body instructs me as to my possibilities and limitations, I begin to understand the sort of creature that I am in relation to all the other creatures, as well as the sort of creature that I am uniquely. Or, as Atkins argues,

> [w]hen I hold something in my hand I not only feel the texture and temperature of the object, but I gauge those qualities *against* my own skin; I feel the frailty of an eggshell against the musculature of my hand, the coolness of ice against the warmth of my blood, the sharpness of glass against the fleshiness of my palm. (Atkins, 2000, p. 336)

One only needs to consider what happens when such bodily perspective is placed in jeopardy by a disease, as in the case of Oliver Sacks "The Disembodied Lady" (Sacks, 1985). Before a routine surgery to remove her gallbladder, the patient, Christina, became unable to stand (unless she looked

directly at her feet), or to maneuver her limbs, and was increasingly incapable of controlling or feeling her body. Greatly distressed, she reported herself to be "disembodied" (Sacks, 1985). Her condition, a loss of proprioception, which allows us to feel our bodies as properly ours -- to be connected to them, and in charge of their actions -- was eventually diagnosed as a rare case of sensory polyneuropathy (Sacks, 1985). Christina's loss of proprioception made impossible exactly those sensations and actions that are both bound, and made possible by, my corporeal existence – by the sensation of the ongoing reality of my particular embodiment (even though the facts or the qualities of this embodiment may change throughout my lifetime). Thus, my embodiment is not just evidenced by my sensations of the world through my body, but also through a recognition of my abilities and limitations. As Atkins rightly notes, it is the body that "allow[s] me to conceive of myself...as both a thing in the world and as a subject of the world" (Atkins, 2000, p. 336).

Yet understanding oneself as an embodied self is not a monologue, nor is it unidirectional. My body is not merely the way through which I experience and interpret the world – it is through this body that the outside world views, interprets, and tells me what it sees. As Lindemann, Walker, Schechtman, Atkins, and I have argued elsewhere, the embodied self narrates, but not just to itself -- it tells stories in part to find some validation of its own embodied experiences in the world of embodied others in order to engage with them in what Walker has called shared moral understandings (Schechtman, 1996; Lindemann, 1997; Walker, 1997; Atkins, 2000, Gotlib, 2009). That is, because we are fundamentally social, our sense of who we are as embodied beings also very much depends on the perspectives and attitudes of others, and is thus very much a narrative collaboration between one's stories and those of many proximate and distant others. This intersubjectivity takes the form of many diverse narratives by, and about, me. And it is the possibility of their uptake as credible and valued that has a lot to do with my perceived legitimacy as a credible and valued embodied moral agent. In other words, my embodied self -- who I am, where I belong, what I seem to be able to do (and not do), what duties and rights I might claim, what views others might have of me, and so on -- is shaped out of the first, second, and third person stories that come to define me within shared social, political, and moral spaces.

What this suggests is that despite our first-personal embodied encounter with the world, from our first moments, we are defined, evaluated -- indeed, even created -- by embodied others: We are born through the bodies of

others, we are wounded, praised, loved, and sometimes killed through the bodies of others, and, in the end, we require those others in our final moments. I am viewed by the outside world as a body with long, dark hair, or in a bright dress, or as short, tall, fat, thin, with blue eyes, with brown eyes. Or sometimes, more significantly, one who is brown, black, white, male or female or neither or both. A body that can walk, or cannot. A body that is relatively intact – or is not. That is, I not only perceive, but also am perceived, through my body -- I am seen, felt, heard, and evaluated as a physical presence. And these stories of and about my body become a part of the constitution of my identity. I begin to see myself as short, tall, male, or female not only descriptively but normatively, with all of the social, psychological, and moral baggage that these labels impose. And I am shaped as a moral agent by this seeing, feeling, hearing, touching -- I integrate into my identity all the normative judgments, valuations, rules, rights, and affections that first find their way through my embodied self.

To summarize, embodiment is not only a fact about us as human beings, but also creates the necessary corporeal perspectives from which we are able to participate, well or poorly, in life itself. Indeed, changes within and without my body can alter the nature and quality of my lived experiences: illnesses, accidents, parenthood, different physical and social surroundings, and so on can alter one's horizons while introducing an awareness of some very different features of one's now altered world. And how I attune and re-attune my perceptions of my lived experiences is a matter of narrative creation, of storytelling. As Atkins reminds us, this social, narrative approach to embodied identities

> has been central to much feminist and civil rights scholarship, for example, which has demonstrated some of the ways in which the social significance of the female body has limited the kinds of experiences that women have been able to have, and so has limited the meaningfulness of the worlds and lives which women have lived. The valuing and devaluing of certain bodily capacities and the social practices in which those values are expressed informs one's view of the world at the same time that it forms one's sense of who one is. Those socially articulated values determine what it is like to *be* female, black, handicapped, working class, etc. I become intelligible to myself through the regard of others, through the subject positions I can occupy in social discourse. Therefore, my sense of who I am and what is 'mine' cannot be isolated from my social setting. (Atkins, 2000, p. 341)

Thus, the question of who we are -- our identities -- is not only fundamentally connected to the fact of our embodiment, but also to the stories we, and others, tell about this embodiment, and the kinds of moral agents we become as a result. What I turn to now is an extension of this discussion of embodiment into new narratively-constituted virtual spaces of *Second Life*.

2. Embodiment 2.0

Before proceeding any further, it seems important, both epistemically and for the sake of clarity, to locate myself within the discourse on enhanced virtual embodiment generally, and within *Second Life (SL)* particularly. Even though this project grew out of my ongoing research into issues of identity formation, narrative practices, and the importance of embodiment in moral work, I initially came to *SL* as a curiosity, partially motivated by philosophical worries about virtual identities, partially moved by my interest in fictional universes, and, perhaps most personally, because I wanted to see *what it was like*. What I discovered was not merely beyond what I expected in terms of the seriousness with which participants constituted and regarded their virtually-embodied selves, but also far exceeded the extent to which I imagined myself to be drawn into this new, strange-yet-familiar world. Thus, although I did not approach my time within *SL* as a formal field study, I nevertheless do rely, at least to some extent, on personal experiences, encounters, and reactions in making some of my claims and conclusions, situating them within the theoretical background provided by my previous work on identity, embodiment, and narrative. However, importantly, for the sake of factual rigor and a broader view of *SL* as a phenomenon, I also focus on the more structured (and extended) ethnographic and anthropological data of Boellstorff and McKeon and Wyche, among others.

So what is *Second Life*? Launched in 2003 by Linden Lab, it is intentionally designed as an environment that is to some extent a blank canvas (with a number of rules) to be peopled, imagined, and constructed by its users. Largely inspired by the hyper-reality of the "Metaverse" environments of Neal Stephenson's novel *Snow Crash*, *SL* is intended to echo his radical vision of users who conduct the majority of their lives while fully immersed in its non-physical spaces. Indeed, it was to become exactly that: a place where anyone could create an avatar body, as well as the space that the body occupies, limited mainly by one's imagination and technical abilities (Boellstorff, 2008).

Although more recently, its popularity has somewhat weakened, the users of *SL* have been known to number in the millions. I begin by considering how *SL* resembles a narratively embodied landscape, and why this matters.

2.1 Virtual bodies, virtual spaces

If *SL* is to be understood as a new arena for theorizing human embodiment, perhaps the best way to start is to view it as a reflection of the familiar tasks of embodiment -- communicating, working, setting up a household, dancing, and so on. In many ways, *SL* is very much like any other intersubjective space, where bodies congregate and *do things,* alone or together. The difference, of course, is that these bodies are not flesh-and-bone, and the "spaces" in which they congregate are not spaces in the usual senses of the term. In fact, they might at first seem to be no different from the characters in video games -- pixelated heroes and heroines, dressed in often skimpy or otherwise revealing clothing; or else creatures who do not share any of our human corporeal shapes at all. But the pixelation and appearance, I suggest, is where the similarity to video games often ends.

Generalizing a bit to the more advanced, later generations of video games, while playing, one very consciously and deliberately participates in two kinds of practices: (1) the aesthetic practices of creating one's avatar; and, more importantly, (2) the instrumental practices of using one's avatar for a largely singular purpose: to succeed at whatever objectives the game requires. The rules and the functions of the game-avatar are strictly set, and one (generally) cannot engage in activities that have little to do with the *telos* of a given game (or one does so at one's peril, resulting in, for example, being "killed" or otherwise eliminated). And while one can certainly project oneself into the game-avatar and see him or her as the continuation of one's self, that projection is usually limited in scope: one is a warrior for only as long as the virtual battle lasts, at which point both the simulation and the projection are over. That is, the user, in her connection to her avatar, is necessarily limited by the rules governing the game -- she cannot engage her creation in activities that might be meaningful to her if they lie beyond the moves prescribed by the game's design. Or, put another way, one is limited by the stories already written by others -- one is necessarily guided by master narratives that enforce character, teleology, and generally frame one's deliberative spaces in a way that

limits both what the deliberation can be about, and how the products of that deliberation might be enacted on the screen.

By contrast, *SL* has no teleology that is enforced from the outside. Certain in-world rules of conduct notwithstanding, participants are not bound by a third-personal notion of what it means to "win," how to "play," or what limits to place on one's avatar. In most ways that count, one is on one's own not only in terms of where one spends one's time in this virtual universe, but also how one does so, with whom, and what the experience might mean or signify. Designed on what has come to be known as an open-world, or a "sandbox," model, individuals within the *SL* universe are free to create, explore, and interact with each other, and with their environments, in ways that were previously reserved solely for encounters within the physical world. Like children in a sandbox who are limited mostly by their own imaginations, abilities, and desires, *SL* participants are not constrained by pre-existing objectives or storylines of its developers. What this suggests is that the virtual landscape of *SL,* unlike a video game, with its linear, instrumental story-lines and goals, much more resembles our non-linear, complex, and fairly open-ended corporeal environment. But perhaps the simplest way to think about *SL* is this: it is much less like a choose-your-own-adventure book in which the reader is asked to select among several possible pre-written story-lines, and thus endings, and much more like a book with blank pages and colored pencils with which I can write the story of who I am – in which I can continue creating an identity, rather than enact a role or complete a task wholly authored by another.

These new identities, one's avatars, similarly have no pre-established limits, save for those imposed by the technology itself. Whether on a whim or as a result of deliberate planning, one can switch between species or create hybrids, consisting of human beings, animals, plants, and even inanimate objects. In the absence of meta-narratives, the stories to be told with and through one's avatars (one can have multiple embodiments) can be as quotidian as setting up, and running, a virtual household, or number among those much less ordinary, stretching one's non-virtual self intellectually, emotionally, sexually, and yes -- physically. To what extent participants actually engage in the more adventurous embodiments will be addressed a bit later, but for now, what is at issue is whether the possibilities for such bodily imaginaries, and the subsequent radical narratives of the self, are present within *SL.* Indeed, they are.

Given the above, I can now make two claims about *SL:* First, the relationships of *SL* users to their virtual spaces and avatars are different from other non-physical encounters not just in degree, but in kind. Rather than allowing the user to participate in a (however complex) pre-established set of meanings and possibilities (for action, for identity creation, for ways of being), genuine practices of narrative meaning-making emerge that, due to this open-endedness, begin to cross the boundaries between the corporeal and the virtual. Thus, the virtual world of *SL* in important ways reflects the corporeal, distinctly *narrative* spaces of the physical universe, where individuals and groups come to create, witness, and morally engage with each other's stories.

Second, these practices of creating and manipulating one's avatar in virtual spaces open up new possibilities for what embodiment itself means by challenging many of our established norms of bodily identity and physical constitution in two distinct ways: first, by *extending* the self into a virtual domain; second, by *transcending* the physical self through a narrative engagement with certain aspects of one's identity that have remained hidden, that were previously unknown to the participant herself, or the expression of which was unsafe in physical spaces. I now turn to the questions of how *SL*'s participants engage in these practices, and why they are so significant for the development of their identities.

2.2 The avatar and its possibilities

One's avatar -- its selection, design, and manipulation -- is central to both extending and transcending of the physical self. "Avatar," derived from the Sanskrit *avatara*, suggests "the idea of a kind of transubstantiation, the incarnation of life in a different form" (Tofts, 2003, p. 56). I take this etymology to be significant, not merely because it differentiates the making of one's avatar from merely temporarily borrowing a character for a limited purpose, but because it is suggestive of the possibility that the act of avatar creation has something to do with fundamental changes in one's identity-constituting stories. The thought goes something like this: In engaging in the practices of making an avatar, we both transfer a part of ourselves to this entity, but also experience an expansion in the definition of who we take ourselves to be, partly as a result of placing ourselves in this new "skin."

This "transubstantiation" takes place in the following manner: The process of creating an avatar begins with selecting a name (along with a "birth date,"

the only element that cannot be altered). Although in its initial form, the avatar is a rather "basic" male-or-female body, its complexity increases with nearly endless customizability. In its finished form, an avatar can appear as a human male, female, neither, or both; or child; or, indeed, as an animal, a magical creature, or even an inanimate object, such as a spaceship or a boat. What one does with one's avatar is really up to its creator: without a pre-existing narrative of either self or space, authorship can be as simple or as complex as one desires. In fact, anthropologist Tom Boellstorff suggests that in the world of *Second Life*, embodiment could be almost anything:

> A man spends his days as a tiny chipmunk, elf, or voluptuous woman. Another lives as a child and two other persons agree to be his virtual parents. Two "real"-life sisters living hundreds of miles apart meet every day to play games together or shop for new shoes for their avatars. The person making the shoes has quit his "real"-life job because he is making over five thousand U.S. dollars a month from the sale of virtual clothing. A group of Christians pray together at a church; nearby another group of persons engages in a virtual orgy...Not far away a newsstand provides copies of a virtual newspaper with ten reporters on staff; it includes advertisements for a "real"-world car company, a virtual university offering classes, a fishing tournament, and a spaceflight museum with replicas of rockets and satellites. (Boellstorff, 2008, p. 17)

An individual can also have a number of avatars (called "alts"), all with different names, and very often with different enough characteristics that it becomes a challenge (if not an impossible task) to figure out if different alts belong to the same user. This emphasis on authorship, rather than on skillful direction-following or clever manipulation of set choices, can be understood as a shifting away from role-playing, and a move toward "being oneself", or extending oneself, in virtual space (Boellstorff, 2008, p. 117). Indeed, *SL* offers an experience that is immediately, and intimately, embodied – one that is not a weak simulation of corporeal life, but a continuation of it in another guise that at times proves to be a more accurate reflection of one's subjective identity:

> Residents, for instance, might say that a particular animated chair caused their avatar to sit in an "unnatural" manner in comparison to the more "natural" animation they typically used. Virtual embodiment could even be understood as more authentic than actual-world embodiment; as one *Second Life* resident put it, "this is how I see myself on the inside." (Boellstorff, 2008, p. 134)

Using this unique virtual body (and communicating either through spoken or typed speech), one can engage in a number of identity-constituting activities: (a) Casual communication, where one might converse with other individuals as a part of "interest groups," such as film buffs, political enthusiasts, and yes, even those interested in philosophy; (b) Commerce, involving the making, selling, advertising, and buying of property, clothing, and so on for use in *Second Life,* but also extending the SL economy, with its currency, the Linden Dollars (L$), into the corporeal world by allowing participants to transfer their profits between the *SL* economy and their corporeal bank in "meatspace"; (c) Education: Several institutions, such as colleges, libraries, and government entities, use *Second Life* as a platform either to conduct research, or sometime, offer courses or information about their activities; (d) Music and the arts: Plays are performed by avatars, and artists display (and sell) their physical and virtual works (some of which would not be possible in the corporeal world, whether due to physical limitations, costs, or both) while musicians perform "live" concerts; (e) Virtual workplaces: Companies take advantage of the virtual spaces for meetings or training programs.

SL is also the forum for a number of religious groups and traditions, including a not insignificant representation of "alternative" beliefs and practices. This is not to mention the numerous islands, clubs, stores, restaurants, bars, and "special interest" areas, as well as the thousands of private residences that populate *SL.* By remaining largely grounded in the kinds of activities that make up large portions of our corporeal lives, *SL* extends our "first world" embodied selves into the virtually embodied domain, allowing them to connect emotionally, as during a virtual funeral for a corporeal person, or else through new mixed corporeal-virtual spaces for art and commerce, as during an artist's exhibit of his gravity-defying works to be sold for real dollars, but able to be displayed only in the owner's *SL* home.

This "realness" of *SL* is not only measured by the kinds of activities within it, but by self-reported accounts of what it is like to be a participant. Indeed, individuals reported that "in world," they feel more like "themselves," more like their "inner selves" then in the physical world. An in-world resident noted that in *Second Life*, "you can be who you are, not your [actual-world] body" (Boellstorff, 2008, p. 135-136). Although a number of participants tailor their avatars as closely to their physical appearance as possible and, if we are to trust self-reporting, strive for both physical and personality consistency

between their physical and virtual selves, an even larger number do not. What this tendency suggests is that how one's identity is expanded within the virtual environment is not merely a matter of combining one's physical and virtual selves. Yet, as I have come to understand, for many, neither is it just a matter of creating a virtual persona for the sheer joy of experimentation with the new and the unfamiliar. In fact, the process of creating and living through an avatar becomes a kind of transcendence of one's physically-embodied self. This transcendence is the freedom to be otherwise, and to experiment with various embodiments, stories, and identities that are not possible, or are *viewed* as not possible or taboo, in the physical world. In other words, what I am calling "transcendence" here can be understood as a counterstory -- or counterstories -- "written" by *Second Life* participants through the practices of avatar creation and enactment. And the possibility of such transcendence, however conceived, presents a formidable challenge to the notions of the virtual as insignificant or morally-ephemeral.[1]

The idea that the virtual is an unserious use of the participant's time, and thus largely uninteresting to outside observers, is not new. In addressing some of the related worries surrounding the epistemic "realness" and the moral authenticity of the *SL* experience, Marya Schechtman, relying in part on the findings of ethnographer Annette Markham, suggests that rather than serving as a mere game or an inconsequential activity that is peripheral to one's core concerns, virtual embodiment is exactly the opposite -- it is, in fact, a new kind of identity-constituting *reality:*

> Markham reports that she "found reason to destabilize a traditional idea that the experience of reality is grounded in the physical, embodied world." To her surprise, she says, the residents she engaged with the question "What is really real?" told her "this question was of little relevance to them; rather everything that is experienced is real...[and] our virtual relationships are just as real as our rl [real life] ones." (Schechtman, 2012, p. 331)

---

[1] In fact, Raffaele Rodogno has argued that "online activities may, in different ways, affect our offline personal identity [...] [T]he more important online activities become [...] the more we can suspect that any self-narrative [one] would recount would include events that occur within them[...] [O]ur interpretation of ourselves is constitutive of who we are," and thus our "identity will significantly be constituted by[...]online life..." *See* Rodogno, R. (2012). Personal Identity Online. *Philosophy & Technology.* 25(3), p. 325-326.

Thus experiences in virtual worlds, Schechtman argues, might very well constitute a phenomenologically "genuine" and morally powerful part of participants' lives, "as real as any others" (Schechtman, 2012, p. 332). Rather than just offering a transient instance of play for a physically-embodied being in a make-believe virtual environment, *SL* seems to offer a chance to re-conceive and embody oneself differently in ways that count -- and that last -- for the participant psychologically, socially, and morally. "Virtual reality" thus experienced is virtual only in its presentation. Everything else about it *is just our life*, albeit from startlingly different perspectives.[2]

This stronger reading of the realness of virtual embodiment supports the kind of transcendence claim that I am making here by putting significant pressures on the master narratives of what it means to be a body and a self who is female, male, straight, gay, white, nonwhite, able-bodied, disabled, and so on. And this is the case not simply because one can "create" an avatar whose appearance challenges the prescriptions of social master narratives regarding how one ought to act and appear, but because, if one wishes, one is able to enact, and *reify*, a "forbidden," or hidden, self. Or else one might find oneself surrounded by individuals never (or not often) encountered in corporeal life -- including not only imaginary creatures and fantasy characters, but importantly, individuals whose embodied races, abilities, and backgrounds might very well have been on the periphery of one's everyday experiences. In all these ways, I suggest, one transcends one's corporeal self through a narrative engagement with, and within, a virtual world.

So what does transcendence look like in *Second Life*? Pavia, a female avatar, notes that

> ...I'm not the person you have gotten to know. But at the same time I am. I'm a man in real life, but about three weeks ago I learned that I'm transsexual. I've pretty much known that I was different all my life[...]Here in Second Life I created something new in myself that I never realized was there before. At first it was just role playing, but then I grew to love Pavia. I kept infusing myself into her, but then something unexpected started to happen: Pavia started coming out in the

---

[2]In fact, Nick Yee has argued that while the effects of avatars on participants within virtual worlds have often been explained in terms of the Proteus Effect -- a phenomenon where an individual's virtual in-world identity is dependent on the avatar under which she or he operates – a deeper investigation suggests that it is in fact the Proteus Paradox which results. The Proteus Paradox suggests that "...where we think we are fully in control, unique psychological levers in virtual worlds (such as our avatars) powerfully change how we think and behave" (in the physical world). *See* Yee, N. (2014). *The Proteus Paradox*. New Haven: Yale University Press. p.5.

real world. I became her, she became me. (Boellstorff, 2008, p. 138)

Another avatar who was quite shy in the corporeal world, observed that

> I noted yesterday that I had no problem talking to a complete stranger at the shopping center, simply because I have spent a lot of time in SL recently doing the same thing. (Boellstorff, 2008, p. 121)

Participants also experiment with gender fluidity, where corporeal-world males will not only appear as women (and women as men), but as transgendered individuals, individuals of indeterminate gender expression, individuals whose gender expression changes weekly or daily, and so on.

Why does this avatar fluidity matter? Sherry Turkle has noted that this "gender swapping" makes it possible for people to experience, rather than merely observe, what it feels like to be the opposite gender or to have no gender at all (Turkle 1997). Thus, rather than play-acting or imagining *what it might be like* to be otherwise, participants, by placing themselves in the middle of actual, rather than imagined, contexts, are subject to the reactions, behaviors, attitudes, and moral judgments of others: they do not simply imagine what it might be like to be a transsex woman in a bar – in an important sense, they *become* her. A participant notes:

> I haven't been able to spend as much time getting my appearance to be more what I want which would be more androgynous or possible confuseable with a female. Basically what I feel like I would be on the inside if I had my choice would be female but possibly confuseable with male occasionally- more towards the androgynous female side of the spectrum. SL was really important to me because it gave me the chance to actually try out what I would want to look like if I had a chance to sort of express the transgendered feeling as my actual appearance and you know to be able to look like.

> In my first life, if this were a perfect world I'd try and represent more of my personality in the real world. If I did do this it would meet with a lot of criticism, it's easier to pass himself off as a generic guy rather than a girl. In a virtual world, I find myself acting more feminine, I can't represent much of his personality in the real world

> I live in the rural south, and even though I love it here, people are too closed-minded. On SL I can look like RuPaul and nobody cares. I can be, be more like myself. (McKeon and Wyche, 2005)

Thus, this gender fluidity, for all of its playful, often masquerade-like appearance, can be quite profound, allowing the participant not only to see, but to be seen, as a unique and norm-challenging self. But the perspectival changes that *SL* makes possible are not limited to self-expression and perception alone. Indeed, to complicate the picture a bit more, the third-party, often heteronormative participants who, in a non-virtual environment, might assume the sociocultural privilege of judgment—or of outright discrimination -- are limited within *SL* both by circumstances and by prevailing attitudes. For instance, even given the persistent prejudice against LGBTQ communities, what one begins to see in-world is a significant shift in the general narrative competence of the participants -- not necessarily in the sense of becoming more nuanced interpreters of the plight of others, but in the sense of acquiring an openness, learned through numerous narrative interactions in *SL,* to the possibility of the wrongness of their beliefs.

How this takes place varies. Generally, because a not insignificant number of participants assume non-heteronormative identities, whether to express their more authentic selves or else to experiment in a kind of world-traveling that is not (safely) available within the physical environment, heteronormativity, and the related oppressions of all those who are "other," becomes among some the exception, rather than the rule (Lugones, 1989). Indeed, not only are participants free to engage in what they take to be expressions of their genuine (or experimental) selves, those who previously assumed the relative acceptance within the corporeal world of their explicit and implicit oppression of others face numerous, and challenging, counter-narratives within *SL*'s environments. In fact, discrimination and acceptance often switch places, with the former relegated to the background, while the latter assumes the foreground, peopled with an increasingly diverse looking, sounding, and acting set of avatars whose presence attests to the power of these virtually-embodied narratives of inclusion.

The widening of the boundaries of normative embodiment within virtual worlds is also quite significant for individuals who are otherwise limited by a corporeal sociopolitical environment that too often refuses to accommodate differences. In the case of non-neurotypical participants, Boellstorff observes that

> Second Life's reliance on textual chat instead of voice during the period of my fieldwork, the limited capacity for avatar facial expression, and a general tolerance for delayed or unexpected responses...made it possible for many

residents with autism to be competent social actors to a significantly greater degree than in the actual world. Even residents with what were typically seen to be more minor psychological disabilities, like Attention Deficit Disorder, often found that Second Life enabled new forms of selfhood...in Second Life they were perceived like any other resident...In cases of severe psychological disability, Second Life could enable significantly new forms of selfhood...Joseph...suffered from debilitation schizophrenia...in the actual world Joseph "is a recluse and rarely communicates." But in Second Life "Joseph...explore(s) places for hours, spend(s) time talking, create(s) things. It is amazing to watch him do thing in here that he could never do in RL. For instance, in RL he lives with [his] mother for obvious reasons. Soon after he came into SL, Joseph put a cabin on a piece of land and decorated it with a few small items...[H]e said "This must be what it feels like to move into a college dorm for the first time." (Boellstorff, 2008, p. 147-148)

Moreover, those with physical disabilities find within *Second Life* a unique kind of embodied freedom: Disabled residents of *SL*, no matter how they represent themselves in-world, can walk, run, swim, and even fly. As some participants have noted, they are able to not be "body bound," and engage in the kinds of movement, connection, and communication that they take to be more authentically their own (Boellstorff, 2008). In other words, what virtual embodiment offers is exactly what these participants could not experience in their corporeal existence: to be otherwise, to transcend the limitations of a society that still has not sufficiently embraced accessibility as a moral norm, and to offer counter-narratives about what it means to be a (dis)abled self, challenging the dominant master narratives about what disabled individuals can do, and what they can be (Boellstorff, 2008, p. 137). I suggest that in so doing, they become more fully, more completely, themselves.

Because of my emphasis on the transformative effects of *SL*, it is important to distinguish the claim that the (dis)abled can become more fully themselves within *SL*'s virtual worlds from the claim that *SL* is a kind of an enhancement technology – a super-prosthetic of sorts. While enhancements, whether physical or intellectual, are designed to be a means to be "better than" -- to be a means of transcendence of our physical and psychological limitations -- the sort of transcendence I have in mind here is not one of overcoming the human. Indeed, it is in an important way quite the opposite: It is the embracing of the expansion our experience and understanding of the normatively human -- the self that is more fully *me*,

rather than the superhuman *better-than-me*. What is more, it is about becoming visible to oneself and to others as someone with an identity that is more complex and richer than could ever have been imagined – one is revealed as a multifaceted, rather than merely enhanced, moral agent.

Moreover, my claim about the possibilities for disabled individuals within virtual worlds has less to do with the broad discursive spaces that *Second Life* makes possible, and more with its normalization of the disabled body as an actively agential one. After all, if all *SL* provided was the opportunity for a disabled individual to "fly" via his or her avatar – and even if the individual did in fact view the experience as liberating – it is unclear at best how the predominant master narratives of the "badness" of disability might be in this way challenged. In fact, they might arguably be reinforced: *SL* would merely represent a momentary escape from an ostensibly undesirable (and fixed) identity of being less-than. Instead, something very different is possible: disabled individuals can choose to create a disabled avatar, and, because this avatar neither experiences any corporeal accessibility issues nor need appear non-disabled to avoid such experiences in-world, he or she can subsequently become a part of any community or environment, at the same time normalizing the presence of disabled bodies in able-bodied-favoring locations or environments. The disabled participant is not merely integrated into a robustly diverse online community, but also, through the practices of avatar-making and engagement, enacts an identity – tells a story – that directly challenges damaging presuppositions not only about ability and access, but about the notions of disability itself.

3. A Few (Serious) Worries

Of course, *Second Life*, and other virtual worlds, are not without serious concerns. Specifically, worries about how the actual behavior of a worrying number of participants suggests not liberation, but a furthering of stereotypes -- less a transcendence of worldly limitations, and more a replication of old, persisting hatreds, prejudices, and fears. The more obvious examples are the choices of the avatar embodiments themselves -- young, mostly white, thin or muscular, able-bodied, often (especially in the case of female avatars) skimpily dressed reflections of the media-created ideals.

One participant describes his SL avatar by noting

> He has blonde hair, he's like your stereotypical gay guy, like the whole six-pack abs and stuff, he's basically what everybody wants to look like in our

community. He wears mainly dark clothing, he's outspoken, however he knows when he needs to shut up. I designed him this way because, in a way, its what I hoped to look like in first life.

Most guys just like wearing jeans and a t-shirt, something that shows off their muscle bodies, because they don't have those in real life but they got them in here.

She looks likes she is in her twenties, but 90% of the women I know are in their 40's but all of our avvies look so young, so nice. (McKeon and Wyche, 2005)

In my own, albeit anecdotal, experiences, I have often been asked by other avatars why I (as a female avatar) choose to dress in such a "boring" way -- black jeans, t-shirt, black jacket, fairly flat shoes. I was often offered free outfits that would make me more "exciting" or "attractive," or "fun" – mostly micro-mini skirts, bustiers, ridiculously high heels, see-through outfits, and so on. When I politely refused the offered "appearances," some participants remarked that I was missing the point of *Second Life* – to be the ideal version of oneself, the version that would, finally, erase one's imperfections, flaws, and other "real life" limitations.

Indeed, a

significant numbers of participants that present as female in-world [r]egarded their bust size as a primary concern when creating a Second Life avatar. 'At first I played with an avatar that I thought represented me physically....But not many people talked to me. Now [with a large-chested avatar] people go out of their way to IM me and send me friend requests.' (Burns, 2009)

In fact, in one study of gender representation within *Second Life*, four out of five women interviewed claimed that while an avatar represented their desired appearance or personality, "what was especially striking is how many women, when prompted, said their avatars were "better" than their real selves. Not just skinnier or sexier, but better" (*Ibid.*). The researchers note that

countless Second Life residents are so enveloped in a popular definition of "attractive" that they need no coercion to create a sexually idealized character. In fact, the creation of the sexually-idealized character at the expense of a character more in line with many players' tastes is mostly deemed necessary for making friends. (*Ibid.*)

What these accounts seem to suggest is that it would not be altogether unreasonable to view the virtual world's influences on gender representations

and gender politics as a revenge effect of a technology that once held the promise of extending, rather than shrink-wrapping, our notions of embodied identity (Tenner, 1996).   After all, a virtual world that allows, indeed encourages, the extremes might very well blithely look past exploitative gender-based master narratives in favor of yet another technological enhancement.   And if the largely male-dominated virtual worlds facilitate opportunities for the idealization and fetishization of certain physical traits while marketing these experiences as truly novel, liberating ones, then perhaps we ought not be too surprised when, once we move beyond the novelty of technological trappings, we find high-tech versions of the same oppressive narratives of beauty, desirability, and perfection. In part, this might very well be due to the newness of the technology itself, and the often slow pace at which we adapt to its unprecedented freedoms. Yet the worry is that these technologies, regardless of the length of our exposure to them, will simply further habituate us to the oppressive norms of gendered embodiments – except this time, under the banner of the "freedoms" of the virtual. And thus instead of opening up new possibilities for transcending corporeal limitations and challenging narratives of oppression, our virtual lives will merely offer us yet another way to damage ourselves and each other.

Finally, I turn to perhaps the most troubling issue within *Second Life* – that of race. Of course, while the subject of race within virtual worlds very much deserves its own treatment, here, I will just note a few particularly disturbing trends and tendencies.

Although the promise of greater racial variability and acceptance is something toward which a number of *Second Life* participants strive, generally speaking, a large number of avatars tend toward white, or at least light-skinned, embodiments. In fact, as Boellstorff noted

> some residents who tried wearing nonwhite skins reported racist responses, including friends who stopped answering ims and statements that nonwhite persons were invading Second Life. It is not surprising that some residents who were nonwhite in the actual world engaged in forms of racial passing, so that at least one of their avatar embodiments was white. (Boellstorff, 2008, p. 145)

Thus, out of the many kinds of racisms on display in online environments (anything from casual racist remarks, to voice-based racism, focusing on how someone "sounds," to racism based on how "ethnic" avatar handles or actual names appear), *Second Life*, as a highly visual environment, often becomes a stage for the most common of all prejudices -- one based on simple appearance,

or, in *SL* parlance, on one's "skin."   And this is when things become complicated: On the one hand, the freedoms offered by *SL* allow the participants to experience themselves, and to be experienced by others, as new and often challenging identities. Yet on the other, we face two significant moral worries about the consequences of these freedoms as it concerns matters of race.

First, we are faced with the worry about unintended consequences of implicit racism associated with a hardening of group privilege. As Lisa Nakamura argues, participants in virtual worlds who appear as different races, genders, and so on engage in what she calls "identity tourism" -- a kind of a noncommittal, nonthreatening journeying of the already normatively-privileged through the experience of being a non-privileged "other." Indeed, noting the fetishization of Asian characters as an example of just such "tourism", she suggests that

> players who choose to perform this type of racial play are almost always white, and their appropriation of stereotyped male Asiatic samurai figures allows them to indulge in a dream of crossing over racial boundaries temporarily and recreationally...Tourism is a particularly apt metaphor to describe the activity of racial identity appropriation, or "passing" in cyberspace. The activity of "surfing," (an activity already associated with tourism in the mind of most Americans) the Internet not only reinforces the idea that cyberspace is not only a place where travel and mobility are featured attractions, but also figures it as a form of travel which is inherently recreational, exotic, and exciting, like surfing...[I]dentity tourism...allows a player to appropriate an Asian racial identity without any of the risks associated with being a racial minority in real life. (Nakamura, 2000)

Thus, what is often missing from such an experience are the actual consequences of being non-white -- or gay, female, disabled, and so on. One simply tastes, but does not inhabit, the identities that too often come with a social, emotional, economic, and political price. And because one does not have to pay this price -- because one can simply slip in and out of one's "skin" and retreat to the safety of maleness, whiteness, heteronormativity — the act of trying out other identities like so many pairs of shoes threatens to do very little, if anything at all, for either deepening or broadening one's moral universe or enhancing one's capacity to engage in more inclusive practices of moral understandings. In so many ways, one might say that it underscores the in-

group privilege by treating the out-group's identities as merely so much costuming and make-up.

Second, there is, unsurprisingly, the worry about the persistence of explicit, outright racism within virtual worlds. One well-known incident involved Erika Thereian, an *SL* member "trying on" the skin of a photorealistic African-American female as a favor to a friend who created such "skins" in-world. Thereian, ordinarily presenting herself as a buxom blonde, recalled that once she teleported into a region, "...a couple people [are] standing around. One said, 'Look at the n***** b****.' Another said 'Great, they are gonna invade *SL* now.' [...]" (Au, 2008, p. 72-74). In fact, As Wagner James Au notes, such a reaction is not at all rare, and a number of African-American *SL* participants who wish to represent themselves as close to their physical-world appearance as possible, engage in "virtual skin lightening," choosing a skin that "passes" for "Latino/a" (rather than black) as a way to lessen the effects of obvious, as well as more subtle, discrimination (Au, 2008; Peterson, 2011). But we ought not be surprised, Nakamura suggests, for

> [r]ace doesn't happen because of biology; it happens because of culture. Race (and racism) is something that develops when our culture rewards the persecution of a smaller group. Unfortunately, it seems that as our lives move more and more into the digital world, we are migrating more and more of the racism in our culture along with us. (Peterson, 2011)

Thus, it seems that even given the new agency-making potentialities of *Second Life,* many are still bound by an enactment of the same agency-denying and identity-limiting stories. The master narratives about darker skin, sexual appeal, and gender normativity seem to have found another home in the virtual world, protected by the anonymity of an avatar, and thus often uncoupled from the possibility of public shame. And so it seems that in too many cases, our avatars neither extend nor transcend, but merely reflect, albeit with an ability to fly. It is my hope that with time, virtual embodiment will not only be recognized as a fruitful subject of theory, but will become the kind of moral and sociopolitical laboratory that allows us to learn more carefully, to communicate more openly, and to engage more fearlessly through a narrative process that broadens and deepens our identities, and those of others.

Conclusion

In this paper, I have argued that if one begins with the assumption that personal identities are both embodied and narratively constituted, the kinds of stories told through virtual embodiments within virtual spaces ought to be taken seriously by moral theorists. Within these virtual worlds, individuals just might be able to claim the kind of authorship of their stories that goes a long way toward challenging longstanding, and damaging, master narratives. They may be able to explore otherness in ways that are de-othering, or they might find more nuanced, less privileged understandings of what it might be like to be otherwise.

These are, of course, all good, important reasons to include the virtual in our understanding of the personal and the moral. And yet, just as I have some worries about the continued damage of stereotype entrenchment and other limitations of virtual embodiment, I am even more concerned about something that at first seems rather abstract and removed from the more immediate concerns about embodied identities. Specifically, I wonder if the invitation to such a great level of control over one's virtual embodiment that the *SL* universe offers leaves users with a false, and perhaps harmful, sense of autonomy and authorship: Embodiment, it tells them, is whatever you want it to be, as is the environment that surrounds you. If the only limitations (outside the few set by Linden Lab) are ones born of one's imagination, the message seems clear: *your personal identity is, fundamentally, your choice.*

If we allow, even for the sake of argument, that virtual identity and its narratives are as powerful an influence as I have suggested they might be on one's overall sense of self, then can it not also be said that in their libertarian emphasis on total agential freedom, they *unhelpfully* contribute to the mythology of total control over who, what, where, and how we are? And might we then not extend this mythology outward, declaring everyone else to be similarly masters of their fate? If we grant this possibility, we face a dilemma: On the one hand, virtual embodiment is indeed a welcome chance to be otherwise; on the other, it is also an environment largely focused on, and experienced through, total authorial control. Put more simply, the freedoms and creative possibilities of the virtual world might allow us to forget that who we are has just as much to do with circumstances, situatedness, or chance as it does with choice -- and perhaps more. And while this movement away from contingency and contextuality and toward unfettered agency might seem like

an empowering possibility, in the end, we ought not lose sight of its dangers. I fear that by embracing such agential control in the virtual world, we may lose sight not only of how communally, relationally, and often accidentally identities are created, but also of the thought that it is a *good thing* that they are. Thus, with the growth and expansion of virtual worlds, it seems that we ought to take seriously not only their power to re-define personal identities, but also note an emerging master narrative about the self as a monological playground of limitless possibilities in our lives, second or otherwise.

REFERENCES

Atkins, K.M. (2000). Personal Identity and the Importance of One's Own Body: A Response to Derek Parfit. *International Journal of Philosophical Studies.* 8(3), 329-349.

Au, W. J. (2008).*The Making of Second Life: Notes from the New World.* New York: Harper Business.

Boellstorff, T. (2008). *Coming of Age in Second Life: An Anthropologist Explores the Virtually Human.* Princeton: Princeton University Press.

Burns, M. (2009). "The Power of Real-World Gender Roles in Second Life", http://www.pixelsandpolicy.com/pixels_and_policy/2009/11/female -avatars.html [Accessed on 11 May, 2012].

Damasio, A. R. (1995). *Descartes' Error: Emotion, Reason, and the Human Brain.* New York: Harper Perennial.

Gotlib, A. (2009). Stories From The Margins: Immigrant Patients, Health Care, and Narrative Medicine. *International Journal of Feminist Approaches to Bioethics.* 2(2), 51-74.

Johnson, M. (2008). What Makes a Body*? Journal of Speculative Philosophy.* 22(3), 159-169.

McKeon, M. & Wyche, S. (2005)."Life Across Boundaries: Design, Identity, and Gender in SL."http://www.mattmckeon.com/portfolio/second-life.pdf. [Accessed 17 May 2012].

Lugones, M. (1989). Playfulness, 'World'-traveling, and loving perception. In A. Garry & M. Pearsall (Eds.), *Women, knowledge and reality: Explorations in feminist philosophy.* Boston: Unwin Hyman.

Merleau-Ponty, M. (1992). *The Phenomenology of Perception.* London and New Jersey:Routledge.

Meyers, D. T. &Jaggar, A. (Eds.) (1997). *Feminists Rethink the Self.* Boulder: Westview Press.

Nakamura, L. (2000). Race In/For Cyberspace: Identity Tourism and Racial Passing on the Internet, Online Version can be obtained from: http://www.humanities.uci.edu/mposter/syllabi/readings/nakamura.html. [Accessed onOctober 31, 2012].

Nelson, H. L. (1997). *Stories and Their Limits: Narrative Approaches to Bioethics.* New York: Routledge.

Nelson, H.L. (2001). *Damaged Identities, Narrative Repair.* Ithaca: Cornell University Press.

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

Peterson, L. (2011). Racism is everywhere. *The Root.* http://www.theroot.com/views/e-racism-sxsw-panel-examines-prejudice-onlinegaming . [Accessed on 30 October, 2012].

Ricoeur, P. (1990). *Time and Narrative, Vol. 3,* trans. Kathleen Blamey. Chicago& London: University of Chicago Press.

Rodogno, R. (2012). Personal Identity Online. *Philosophy & Technology.*25(3), 309-328.

Sacks, O. (1998). *The Man Who Mistook His Wife for a Hat: And Other Clinical Tales.* New York: Touchstone.

Schechtman, M. (1996). *The Constitution of Selves.* Ithaca: Cornell University Press

Schechtmann, M. (2012). The Story of my (Second) Life: Virtual Worlds and Narrative Identity. *Philosophy & Technology.* 25(3), 329-343.

Tenner, E. (1996). *Why Things Bite Back: Technology and the Revenge of Unintended Consequences.* New York: Knopf.

Tofts, D. (2003). Avatars of the Tortoise: Life, Longevity and Simulation. *DigitalCreativity.* 14(1), 54-63.

Turkle, S. (1997).*Life on the Screen.*New York: Touchstone Books.

Walker, M. U. (1997). *Moral Understandings: A Feminist Study in Ethics.*New York: Routledge.

Yee, N. (2014). *The Proteus Paradox.* New Haven: Yale University Press.

Who Should Enhance? Conceptual and Normative Dimensions of Cognitive Enhancement[*]

Filippo Santoni de Sio [†]
f.santonidesio@tudelft.nl

Philip Robichaud [‡]
p.j.robichaud@tudelft.nl

Nicole A. Vincent [§]
me@nicolevincent.net

ABSTRACT

When should humans enhance themselves? We try to answer this question by engaging in a conceptual analysis of the nature of different activities. We think that cognitive enhancement is morally impermissible in some practice-oriented activities, such as some educational activities, when it is the case both that cognitive enhancement would negatively affect the *point* of those activities (i.e. learning *through a certain kind of effort*) and that we have good reasons to value that point. We then argue that cognitive enhancement should be allowed in two groups of cases, namely in practice-oriented activities, such as recreational activities on which little moral value or social import hangs, and in some prominently goal-directed activities, such as high-responsibility professions, the goal of which has significant moral or social value. Finally, we argue that the use of efficacious and relatively safe cognitive enhancers may even be obligatory in those high-responsibility professions under certain *special circumstances*.

Introduction

When should humans enhance themselves? In this paper we try to answer this question by focusing on a particular kind of enhancement: cognitive enhancement. Typical examples of cognitive enhancers are pharmacological substances like methylphenidate and modafinil, which are reported to improve people's performances in terms of wakefulness, attention, concentration and retention of memory, not only when taken by people diagnosed with mental deficits or disorders, but also when taken by healthy subjects (Repantis et al., 2010). Even though the current efficacy of these substances is far from conclusively proven, and their side-effects not well known yet, their use is reportedly quite widespread, especially among students (McCabe et al., 2005; Weyandt et al., 2013). At present, issues of safety represent the most serious argument against off-label use of such substances. However, in a hypothetical but realistic future scenario in which some of these substances are proven to be both effective and safe, new and possibly more complex moral questions will arise.

We take an analytic approach to the issue of the morality of (cognitive) enhancement. Our approach is analytic insofar as we refuse to take a general stance for or against enhancement. We think that different activities and circumstances require different ethical answers on the use of enhancers. In our view, cognitive enhancers should be forbidden in some circumstances and allowed in others. We also argue that cognitive enhancement may even be obligatory in some circumstances. In order to justify these ethical claims, we will first outline a systematic conceptual framework. Whereas many current ethical approaches focus on issues of authenticity and/or fairness, we think that other crucial normative considerations can be made apparent by engaging in a conceptual analysis of the nature of different activities (Santoni de Sio et al., forthcoming).

Our main claim is that in order to determine whether enhancement is forbidden, allowed or obligatory, one must also attend to the metaphysical characteristics of the activity in question, and in particular whether the activity is prominently goal-directed or practice-oriented. We think that cognitive enhancement is morally impermissible in some practice-oriented activities, such as some educational activities, when it is the case both that cognitive enhancement would negatively affect the *point* of those activities (i.e. learning *through a certain kind of effort*) and that we have good reasons to value that

point. We then argue that cognitive enhancement should be allowed in two groups of cases, namely in practice-oriented activities (e.g. non-competitive Sudoku), on which little of moral value or social import hangs, and in some prominently goal-directed activities, such as high-responsibility professions, the goal of which has significant moral or social value. Finally, we argue that the use of efficacious and relatively safe cognitive enhancers may even be obligatory in those high-responsibility professions under certain *special circumstances*.

We think that our reasoning may have a wide interest, as the arguments that we put forward in relation to pharmacological cognitive enhancement may also be applied - maybe with some modification - to other kinds of human enhancement. We thus hold that the structure and methodology we follow constitutes a template for a fruitful ethical discussion in this and in related areas.

1. The Nature Of Activities

In order to make way for an analytic answer to the question on the permissibility of cognitive enhancement in different activities, we will outline the basics of a theory of the nature of human activities. All human activities are defined through their point[1]. Sometimes the point is an *external* goal (e.g. *gaining* money through one's work as a financial broker), and sometimes the point is the realization of a certain goal *internal* to the practice (e.g. *deploying* certain physical or intellectual abilities in a game, or *enjoying* the company of other people in an informal friendly chat). Admittedly, most human activities are in that respect complex, as they usually have both external and internal points. In other words, human activities have no simple essence. However, either an external or an internal goal is often prominent in the definition of a given activity. On the one hand, certain activities are prominently defined through their external goals - financial intermediation is mainly about gaining money, medicine is about healing people, the military is concerned with defending a state's territory or other strategic interests of a country. In contrast, other activities are prominently defined through their internal goals – friendly chatting is about spending time with friends or acquaintances (rather than merely exchanging information), running is about engaging in a certain

---

[1] See section 2 below. For a more detailed presentation of the relationship between the point and the nature of activities, see Santoni de Sio et al. (forthcoming).

kind of physical activity (rather than merely reaching a given destination faster than walking), reading fiction is about engaging a certain kind of intellectual activity (rather than merely to learn a story).[2] For the rest of the paper we call activities prominently defined through their external goal *goal-directed activities*, and we call activities prominently defined through their internal goal *practice-oriented*. A simple test to realize whether a certain activity is goal-directed or practice-oriented is to try to mentally eliminate either the realization of the internal or external goals of a given activity, and see which one would result in the loss of that activity's point. Would it make sense, for instance, to go out with friends if one did not enjoy their company, or to play a certain game if one did not find the activity amusing or challenging or interesting? As the answer to both questions is negative (setting aside other goal like wishing to develop the friendship or to acquire an appreciation for the games), one may conclude that those are practice-oriented activities. In contrast, would it make sense to work full-time at a brokerage without the prospect of making money or to work as a physician without the prospect of healing patients? As the answer to both of these questions is negative (assuming one does not, for instance, view these mainly as the realization of a childhood dream which it is important to pursue or a promise which has to be maintained, i.e. that becoming a broker or a physician was your ambition or something you promised to do), one may conclude that these are goal-directed activities.[3]

This analysis of the nature of activities has an important implication for the status of the *rules* that apply to them. According to what has become a commonplace in philosophy, rules come in two kinds. Some rules are merely *regulative*, as they set standards for practices that exist apart from the rules.

---

[2] This point has been fruitfully elaborated in the philosophical literature through the idea of internal as opposed to external goods pursued by or through an activity (MacIntyre, 1985). See Schermer (2008) for a discussion of the relevance of internal goods in the debate on enhancement.

[3] Yet again, we are well aware that, as our qualifying comments (in brackets) highlight, in real life activities may often have a mixture of many practice-oriented and goal-directed facets. For instance, as academics we are blessed for having jobs that involve scratching intellectual itches which we might have done (and we sometimes do) even without being paid a salary to scratch. Nevertheless, earning an income is also part of the reason why we do what we do. In real life the situation will always be complex, but we also think that some activities have features in virtue of which they are either significantly more practice-oriented or goal-directed, and this is the simplifying assumption under which we now proceed.

Other rules are *constitutive*, as they are necessary preconditions for the existence of the practices to which they apply (Rawls, 1955; Searle, 1995).[4]

As goal-directed activities are defined through the goal that they attain (the outcomes they bring about), there are no conceptual, or *a priori* limits to *the way* in which or the means by which these activities may be performed, and so the current rules of the practice can be changed without any particular concern for the point of the activity being lost or the nature of it being distorted. Examples of such goal-directed activities would be professions like surgery, civil aviation, and the military, but also scientific research. Here, as long as people are healed, or safely and efficiently brought to their destination; as long as national strategic assets are protected; as long as scientific breakthroughs are achieved, the point of surgery, civil aviation, military, or science is realized. Of course, a violation of or a change in the current rules of the practice in a goal-directed activity like surgery, for instance a rule prohibiting the use of a given technique or technology, may raise different kinds of concern. It may raise *prudential* concerns, as the change or violation of a prudential rule may lead to unpredictable outcomes; but also *ethical* concerns, as the change or violation of such rule may be in conflict with societal values like fairness, safety, or others. However, what is at stake here is the best regulation of an activity, not the nature of the activity being fatally distorted, and its existence thus being jeopardized.

On the other hand, the relevance of internal goals in practice-oriented activities makes certain rules *constitutive* of the activity, constitutive because they are necessary for the existence of that activity, not only for a fair or efficient regulation or perhaps for coordination of it. The presence of constitutive rules thus poses *conceptual*, or *a priori* limits to the way in which a certain activity can be performed. A violation or a change in one of these rules may therefore make that activity lose its point and hence its very nature. Clear examples of this concern come from sport. If one shows up at the starting line

---

[4] Examples of regulative rules are the rules for driving a car or the rules of bon ton for consuming a meal. Examples of constitutive rules are the rules of the game of chess and the rules of language. Whereas it is conceptually possible to drive a car in the absence of traffic rules, and to consume a meal in the absence of any bon ton, it is conceptually impossible to play a game of chess without the rules of chess or to speak a language without any grammatical rules. Chess and language exist in part thanks to their rules. Take the rules away, and the game of chess will not exist anymore (only small wooden pieces being shuffled around on a black-and-white chequered surface will be left). Take the rules of grammar away, and language will not exist anymore (only sounds and doodles on papers). At this level, only a purely naturalistic description of the phenomena remains available.

of a marathon wearing roller-skates[5], they would be not only violating a regulative rule of marathons, but also a constitutive rule of the game. Hence, they would be doing something 'wrong' in *two* different ways. At an *ethical* level, they would be trying to get an unfair advantage against the other runners. At a *conceptual level*, they would be *missing the point* of running a marathon, and hence, given that human activities are defined through their point, they would be simply engaging in a *different kind of activity*. In the latter sense the roller skater would be doing something 'wrong' in a morally neutral but metaphysically loaded sense - wrong in the sense of violating the rules that are constitutive of the relevant activity.

2. When Enhancement Is Forbidden

The debate on the morality of enhancement has been so far particularly hot in relation to sport activities. Even though the use of performance-enhancing substances in sport does not qualify as cognitive enhancement, we will briefly discuss an argument that has been often put forward in the context of this discussion, as this may be helpful to frame an argument against the use of cognitive enhancement in certain educational activities. Many think that the use of performance-enhancing medical substances in sport (often referred to as 'doping') must be banned not only because of the risks for the health of athletes, but also because the use of these substances is against the "spirit" of sport, i.e. it violates the constitutive rules of sport practices. The logic of this argument implicitly rests on our analysis of the nature of activities. In order to fully understand this logic a further conceptual distinction has to be made, that between *coarse-grained* and *fine-grained* descriptions of activities.

As Wittgenstein famously wrote, there is no such thing as a single feature shared by all games, by virtue of which it is possible to decide once for all what should count as 'a game'. Both letting a little rubber ball repeatedly bounce on the wall in front of you while sitting alone at your desk and professional rugby are 'a game'. In a similar vein, it may be said that also a particular game can be played in very different ways in different times, places, and circumstances. There were car races sixty years ago as are there today, and basketball is played by children in parks and courtyards, as it is played in packed NBA arenas. Still, one can draw distinctions among these games by employing coarse- or fine-

---

5 This example, originally presented by Whitehouse et al. (1997), is also discussed in a similar vein by Schermer (2008) and Murray (2008).

grained descriptions of activities. According to a coarse-grained description - one which takes into account only some macroscopic features of the activity - 1950s' and contemporary car racing are *the same game*. Similarly, a courtyard basketball game between children is *as much a basketball game as* an NBA Final. However, according to a more fine-grained description - one that takes into account a larger number of features of the activity, 1950s' and contemporary car racing are *not* the same. When one considers how much cars have changed with respect to their top-speed and their embedded technology, how driving technique has evolved and so on, one will hardly resist the conclusion that today's car racing is *not* the same game as sixty years ago. Similarly, once the marked differences in rules, skill level and training that are present in NBA games are compared to typical courtyard basketball games, it is natural to infer that these are two *different games*.

The possibility and legitimacy of fine-grained descriptions of sport activities may constitute the basis for a moral argument against the use of enhancing substances. Consider a sport S in which the use of a new powerful pharmacological enhancer P is not, as things stand, included in a fine-grained description of the sport's rules and regulations. According to a sufficiently fine-grained description of S, it can be said that athletes using P engage in a different activity, i.e. S_E. As a consequence, *if* there are substantive reasons to value some aspect of traditional sport S that is going to disappear in enhanced sport S_E,[6] *then* there is a *prima facie* moral reason to oppose the use of the substance P in S.[7] In the case of performance-enhancing medical substance in sport, unlike what many people think, it is debatable whether either of the two conditions (the conceptual and the moral) for ban are met.[8]

6 This seems to be, for example, the concern behind the words of President's Council on Bioethics, when he says that the sportsmen who would use biotechnological enhancements would be bad sportsmen–"not simply because they cheated their opponents, but because they also cheated, undermined or corrupted themselves and the very athletic activity in which they seem to excel." (President's Counsel on Bioethics, 2003, p. 161–164), also cited by Schermer, (2008, p. 86).

7 Admittedly, in order to make these prima facie reasons against the use of enhancers conclusive more is required, namely the prima facie reason against must not be outweighed by stronger reason(s) for the use of such enhancers.

8 A discussion of this aspect falls beyond the scope of this paper. See Savulescu et al. (2004), where they convincingly claim that: as for the first requirement, we must not necessarily stick to the Ancient Greek view of sport, according to which only natural talent and strength has to be measured and praised in sport performance; as for the balance with other reasons, given that professional sport cannot be cleaned up (even if we wanted), a regulated use of drugs should be allowed in it, in order to achieve the best protection of athletes health (more control would be possible).

However, these conditions seem to be met in the case of the use of some pharmacological enhancers in some educational activities. Technology has in the past decades dramatically changed intellectual activities. Digital search tools and word processing have represented massive performance enhancers for research, teaching, and study, as they have allowed substantially greater productivity by reducing the time for research, the costs of revision and the speed of editing written texts. Under the fine-grained approach, academics, teachers and students today are therefore clearly engaged in a different kind of research, teaching and study activity than thirty years ago. Something similar may happen if substances, such as methylphenidate (Ritalin), were able to substantially change productivity in research, teaching, and study[9]. How should then the use of pharmacological enhancers be regulated in the intellectual field? According to our framework, the answer to this question depends on the particular nature of the intellectual activity in question. Whereas cognitive enhancement may *not* raise particular conceptual and moral concerns in the case of scientific or academic research (as these can be seen as prominently goal-directed activities)[10], we think that in regards to some educational activities the use of cognitive enhancement *does* raise moral concerns that may constitute ground for forbidding their use. If among the relevant goals pursued by a school/course/class/exam is, for instance, that of teaching how to exert certain intellectual efforts without recurring to any "external" support, or how to cope with certain psychological and motivational challenges without recurring to medicines, or maybe to educate pupils to arrange their work schedules without doing last-minute rushes of study; and if - according to a hypothetical scenario - pharmacological substances would to some relevant extent relieve the students from exerting those efforts and learning the psychological strategies and acquiring the required motivational and organizational capacities; if this is the case, then the use of cognitive enhancement would be *conceptually* problematic, as it would turn the activity into something different. In addition, if we have reasons to value the non-enhanced version of the activity (because, for instance, we think that acquiring those abilities is part of our conception of a good education), then we have a *moral* argument against the use of enhancers in this context.

9 It is controversial whether and to what extent current pharmacological cognitive enhancement is efficacious. For a survey on the scientific literature see Goold & Maslen (2014)
10 Though the case of academic research is very interesting, we are not discussing this here for reasons of space.

In this sense, the normative reasoning on the use of pharmacological cognitive enhancers in various educational activities is arguably similar to that regulating the use of *non*-pharmacological enhancers like books, the Internet or calculators. Should students be allowed, for instance, to open books or access the Internet or using calculators during an exam? We think that the answer to this question depends on what the point of that particular exam (and course) is. If the exam (and the course) aims to test (and foster) the students' ability to *memorize* certain notions or to carry our calculations by hand, then probably students should not be allowed to keep books and internet connections open or to use calculators during the exam (and they should also be encouraged to do the same at least in some steps of their study at home).

Notice that even though it applies mainly to young students, this argument against the use of cognitive enhancers is not directly dependent on the young age of those to whom it applies. In fact, the main point is not that by not being (fully) autonomous agents they may not be left the choice if and when to use medical substances. The point is rather that the *nature of activities* in question may exclude the use of such enhancers. In this sense, the argument may also possibly apply to adults involved in similar activities, even though, as a matter of fact, the kind of educational activity above envisaged is more likely to concern young people.

Also, the moral considerations that do the work in our argument do *not* rest on a concern for fairness or worries that enhancement amounts to cheating. Admittedly, concerns for fairness *may* also be present in some of the circumstances that we presented, i.e. when the acquisition of some limited benefit depends on the results of a given educational activity or exam, and the exam can thus also be seen as a competition for the distribution of those benefits. But this is not the relevant consideration here. In fact, the consideration that we identify would be present and morally relevant also in clearly non-competitive educational activities, such as quizzes taken in introductory college courses. Here, as there is no relevant competitive element, the moral wrongness of enhancing would mainly derive from enhanced agents not realizing the (valuable) point of the activity.

3. When Enhancement Is Permissible

In this section, we turn to a discussion of cases where cognitive enhancement may be morally permissible, that is cases in which individuals must be left free

to decide whether to enhance or not. We think that there are two possible scenarios in which cognitive enhancement is permissible. The first involves all those practice-oriented activities that are not appreciably morally significant. Here, the fact that an agent is now engaged in an enhanced as opposed to an unenhanced activity has no relevant personal or social consequences. Thus, even if the use of cognitive enhancers results in a change in the nature of the activity in which an agent is engaged, there is no moral reason to forbid this change. In such cases, agents should be permitted to utilize cognitive enhancement technologies if they wish. The second way in which cognitive enhancement can be permissible involves strongly goal-directed activities. Here, the benefits of better *results* achieved in the enhanced version of the activity counts as a decisive moral reason for allowing cognitive enhancement.

The first way to identify an activity wherein cognitive enhancement may be permissible is to establish that the enhanced activity in question is morally innocuous. A paradigm class of morally innocuous activities is non-competitive hobbies or passtimes. Such activities are typical practice-oriented activities - the point of engaging in a hobby is given by doing certain things in a certain way. The point of collecting isn't merely to amass a collection. Rather, what makes someone a collector are her various acts of collecting. Although it may be difficult to trace all downstream morally relevant effects of engaging in these activities, there are clearly certain cognitively demanding hobbies that seem in themselves to have no social effects. In such a case, we wouldn't expect the fact that someone engaged in a different, enhanced version of the activity to have any moral significance. Consider an avid Sudoku player who has grown bored with solving puzzles unenhanced and who is interested in seeing how much more quickly she can come to the solution while in an enhanced state. As long as she is not entering into Sudoku competitions or otherwise benefiting from her performances, her activity seems so insulated from the social sphere as to render it a completely morally innocuous enhanced activity. To be sure, since Sudoku playing is a practice-oriented activity, the point of which is using one's wits to solve a puzzle as quickly as one can, according to a fine-grained description of the action, a cognitively enhanced Sudoku player is engaged in a different activity. In this respect she is similar to a cognitively enhanced student. And, we acknowledge that the fact that the enhanced Sudoku player changes the nature of a practice-oriented activity provides at least some reason to suppose both that she is doing something *conceptually* wrong and, thus, that she may have a reason to abstain from enhancing herself. The difference,

though, between the enhanced student and the enhanced Sudoku player, lies in the moral significance of the unenhanced form of the activities. In the case of the Sudoku player, there appears to be no significant *moral* reason not to allow her the freedom to engage in the cognitively enhanced version of her hobby - nothing of moral significance hangs on her solving Sudoku puzzles while unenhanced.[11] This conclusion is buttressed by the fact that there is an independent value in allowing agents considerable liberty to make their life choices. We respect the value of freedom by allowing others to engage in any morally innocuous, enhanced activity they want.

The second way to identify activities where cognitive enhancement may be permitted is simply to attend to strongly goal-directed activities. In this class of cases, the permission to use enhancers doesn't rest on the moral innocuousness of the enhanced version of the activity. Rather it relies on the fact that enhancement fosters the external goal of the activity. Consider agents who have professional obligations to undertake difficult and temporally extended actions or series of actions in a way that benefits or protects other agents under their care or supervision. Prominent examples of such professionals are surgeons and airplane pilots. In both cases, the relevant agents can be understood to be engaging in goal-directed activities. The goal of medical practice in general is typically taken to be the relief of suffering and the cure or treatment of disease, and the goal of civil aviation is to get passengers to their destinations safely and without incident. Given that these professions have such explicit goals, the activities that physicians and pilots are engaged in are not threatened by the use of cognitive enhancers in the way that enhancement technologies threaten sport or educational activities that are practice-oriented. Consider that a surgeon who took a cognitive enhancer before undertaking a long and complex surgery would still be uncontroversially engaged in the activity of 'performing a surgery'. The surgeon is still working toward the goal of her activity. Moreover, the fact that the surgeon utilized some cognitive enhancement technology in order to reduce the likelihood of mistakes does not seem to pose an immediate threat to the traditions, or the "spirit" of the practice. Indeed, one way of understanding the history of medicine is as a series of challenges and changes to traditional medical practice that occur as new techniques and technologies are developed. When cardiologists began to employ MRI technology in order to detect vascular

---

11 We are assuming that the Sudoku player is not participating in competitions or otherwise benefiting from her performances. She's just solving them alone.

problems, they did not cease to practice cardiology, nor did they cease to practice it well. Rather, they not only continued to be engaged in the same goal-directed activity but, by adopting a technology that was more sensitive to the presence of heart problems than existing technologies, they were arguably engaged in a more virtuous instance of this activity. Similarly, cognitive enhancement technology might also allow surgeons to engage in their activities in a way that better fits with the goals and thus the essence of surgery. In this sense, a surgeon who chooses to enhance does not cease to be performing surgery – the nature of her goal-directed activity remains unaffected. There appear to be good reasons to forbid the use of enhancement in this activity.

The same can be said about civil airline pilots. They are also engaged in an activity that has a specific goal, namely that of transporting the passengers safely to their destination. A pilot is expected to use whatever tools or technologies allow her effectively to meet this goal. Consider a long-haul pilot who knows that she can stay more alert for longer periods of time if she takes a cognitive enhancer like modafinil. Because her activity as a pilot is goal-directed, the fact that she engages in an enhanced version of activity does not entail that she is not engaged in the activity of transporting passengers safely. There is a certain similarity between pilots who utilize new radar or navigation technologies and those who choose to utilize some safe, effective cognitive enhancer that is known to reduce the likelihood of fatigue-related mistakes. In each case the new technology offers a different mechanism for securing the goal of providing safe flights and thus to realize the essence of the activity.

The surgeons and pilots we have been discussing are engaged in clearly and strongly goal-directed activities. These goals might be promoted by utilizing cognitive enhancers. Importantly, the realization of these goals is morally significant, and thus it makes sense to judge it morally permissible.

One might object that things are not so simple. By using enhancers, surgeons and pilots may exert competitive pressure on their colleagues to enhance themselves as well even if they would prefer not to. This pressure and the associated loss of freedom count as a moral reasons *against* permitting cognitive enhancement in such professions. Although it is true that this reason has to be factored in the balance, it does not seem to be a decisive one. That we already tend to think that social benefit-based reasons to allow enhancement in professions are stronger than competitive-pressure-based reasons *against* enhancement is shown by the ready acceptance of other more common and widespread ways to gain competitive advantage in professions, (e.g. using

expensive technological devices or having longer working hours). We thus conclude that it would be permissible to allow high-responsibility professionals like surgeons and pilots to use safe and effective enhancement technologies. In the next section we suggest, however, that it may actually be *obligatory* for surgeons and pilots, under special circumstances, to enhance themselves.

4. When Enhancement Is Obligatory

In this section, we introduce and refine a moral principle from which it seems to follow that certain people engaged in particularly high-stakes professions like surgeons and pilots might even have an *obligation* to enhance (Vincent, 2011). After briefly considering some possible objections to the principle, we tentatively conclude that despite these objections, there may be certain cases where cognitive enhancement would be obligatory.

The moral principle that may underlie an obligation for certain professionals to enhance themselves is what we call the Easy and Safe Beneficence principle or ESB.

ESB: If an agent can perform a certain easy, safe, and permissible action A that will allow her to reduce or eliminate suffering for those depending on the agent, then she should A.

There are three things to note about ESB. The first is its close relation to the principle of beneficence that is common currency in professional ethics. ESB is not meant to differ markedly from the principle of beneficence in terms of demandingness or context of application. Second, the easiness of the required action A is explicitly included in order to rule out cases where the beneficent action would be overly costly in terms of effort or self-sacrifice. Indeed, many are probably inclined to think that caretakers should go to significant lengths to protect those under their care. In this respect, ESB is a rather modest principle. The required action here should not be intuitively overly demanding. Third, the safety of the required action is intended to rule out cases where a beneficent action puts the agent at considerable risk of harm. Again, many are no doubt inclined to think that caretakers may frequently be expected to take on rather significant risks to their own well being, but ESB is much more modest. Though we do not have space to discuss the ESB principle at length, we will assume that it is no less plausible than the standard principle of beneficence. Indeed, given that it only requires beneficent actions that are easy and safe, it is likely more plausible than the more general principle. The

questions for the remainder of this section are whether ESB can withstand scrutiny and, if so, whether the existence and use of cognitive enhancement technology could establish an obligation for certain surgeons or pilots to use it. There are several ways one might object to ESB. First, one might note that easy and safe action A may not be the only way of reducing or eliminating suffering to the degree that A does. If in a given situation there were some mechanism for combating fatigue and subsequent fatigue-related errors that did not involve cognitive enhancement, and if this alternative mechanism was itself easy and safe, then ESB would be false. The mere fact that some easy, safe, and permissible action confers benefit is not sufficient for establishing the obligation to perform that action when another (especially better) alternative exists. For example, a surgeon might arrange to have someone take over once fatigue has set in and there is an increased risk of error. And, a pilot might arrange a similar hand-over after a certain period of time. In short, ESB might not entail a duty to enhance because enhancement might simply not be necessary for realizing the reductions in suffering.

Second, ESB might be objectionable on grounds that the action that it requires might be contrary to what might be called Williamsian reasons that involve the agent's personal values or life-projects. In his writings about the virtue of integrity, Bernard Williams famously claimed that such considerations could function as limitations on what morality can require of us. He warns of alienation that agents might suffer if they were required to perform certain actions merely in order to realize some benefit to others (Williams, 1973). On this account, if taking cognitive enhancers were something that ran counter to the deeply-held values of a certain surgeon or pilot, then it would be false that she should take them, despite their benefits. Perhaps such agents take extreme pride in being able to perform their professional duties with their faculties unaltered or unassisted, much in the way that certain mountaineers prefer to forego supplemental oxygen when ascending the world's highest peaks. If surgeons or pilots genuinely identified with this kind of practice-orientation, then there would be integrity-based reasons to think that they could not be obligated to enhance.

Finally, one might object that even if ESB is defensible on philosophical grounds, it simply doesn't apply in the kind of cases under consideration given that cognitive enhancement technologies are not known to be effective in reducing the sort of fatigue-related mistakes that pose a threat to patients and passengers. Though there is evidence that certain fatigue-related loss in

cognitive capacities can be forestalled by using certain pharmacological enhancers, some studies have disputed this (McCabe, Teter, and Boyd, 2004). In addition, they are not universally effective, and they haven't been shown to improve the loss of psycho-motor performance caused by fatigue (Sugden et al., 2012).

In the face of the first objection, ESB must be reformulated. We must stipulate that ESB would only generate an obligation to enhance, when there is no alternative course of action that would realize the beneficial effects. Indeed it is easy to imagine cases where there are no such alternatives. There simply may not be anyone who can take over once fatigue sets in.

The third objection is well-taken. Though it does not directly call ESB into question, it does point to an important difference between new technologies, such as MRI machines and autopilot programs, which have established track records of offering the promised improvements. Many cognitive enhancement drugs, such as modafinil, have not undergone such extensive testing, and so the relevant professionals cannot be sure of the benefits. In addition, even if cognitive enhancement is shown to reduce fatigue related errors in the relevant professionals, the drugs may not affect individuals in the same way. It may be unsafe for some to take it, and it may have either diminished or, worse, no fatigue-diminshing effects. For this reason we will make a concession and a clarifying assumption. The concession is that *at the present time* there is not enough evidence to sustain the claim that ESB entails that surgeons and pilots should undergo cognitive enhancement. The assumption that we will work with is that there is some cognitive enhancement technology that *is* known to be both safe and effective for most if not all of the relevant professionals.

The second objection is not so easily dealt with. In order to assess whether it constitutes a legitimate challenge to the claim that surgeons and pilots have an obligation to take safe and effective enhancements in order to achieve benefits that are unattainable through any other means, we must assess the strength of the integrity-based reasons that agents may have to refrain. It is helpful to compare this response to cognitive enhancement to examples of similar resistance that is based on either on conscientious objections or bald appeal to tradition. As an example of the former, many physicians refuse to perform abortions or offer reproductive counseling on grounds that they conflict with their religious beliefs.[12] As an example of the latter, note that

---

12 For a helpful discussion of this topic, see Savulescu (2006).

physicians were notoriously resistant to accepting aseptic techniques that were clearly supported by the latest science at that time (Gawande, 2012). It is possible that recalcitrant surgeons realized that there was something to be said for the new methods, but nonetheless took it to be permissible to stick with traditional practice. Thus, resistance to cognitive enhancement may be based either religious or moral beliefs, or it could also amount to the claim that 'this is not how it's done!'. In the latter case, we think ESB is on solid ground. By definition, new technologies differ from standard practice, and it would be folly to suspect that the moral reasons *to* enhance are outweighed simply by considerations having to do with the value of tradition. If the goals of the relevant activity would be better achieved through cognitive enhancement, as we are assuming for the sake of argument that they would be, then one might actually argue that the surgeon who refuses to enhance is acting impermissibly. Just as we look back at stubborn attempts to keep traditional practice alive with justified consternation, we should be wary of taking too seriously any attempt to preserve the status quo for its own sake.

Regarding resistance to the ESB principle and thus to the obligation to cognitively enhance that is grounded on deeper moral or religious objections, there are several things to say. The first is that, as with standard cases of conscientious objection, there may be an obligation to refer the patient to another physician who would be willing to realize the benefits of the enhancement technology (Savulescu, 2006, p. 296). Second, it is important to explore what exactly underlies the objection. Two obvious things come to mind, namely that cognitive enhancement technologies function by affecting the physician's body. More specifically, they affect their brain thus the mental states of the professional. Because of this unique mechanism of action, it is understandable that some professionals might think it problematic to be expected to alter their own physical and mental states as a means of realizing some benefit to those under their care. If this consideration is what underlies any resistance to cognitive enhancement, then we think that the best reply is to note that the two professions we are currently considering are already quite physically and mentally demanding. We wonder whether there is a difference in kind between expecting surgeons to endure the taxing physical and psychological work involved in long surgeries or flights and the expectation that they alter their bodies and minds with an enhancement technology. In addition, the mere fact that the mechanism of realizing the benefit involves changes to the agent's physiological states seems far from decisive. Although it

is a fanciful case, imagine there was some way of drastically reducing the likelihood of fatigue-related error by having the surgeon or pilot periodically run on a treadmill (perhaps because the resulting increased blood flow was shown to have fatigue-mitigating effects in the short term). We doubt that many would deny that the relevant professionals should be obligated to do so. Of course, there may be some other integrity-based reason for thinking that certain professionals would have no obligation to cognitively enhance. However, the objections we have considered, which we take to be the most plausible, fail to pose serious problems for ESB. We thus tentatively conclude that there very well could be moral obligations for certain individuals in certain circumstances to cognitively enhance.

Conclusion

According to our nature-of-activity approach, there is no simple answer to the question on the normative treatment of cognitive enhancement. Our reasoning has showed that cognitive enhancement is morally *impermissible* in some practice-oriented activities, such as some educational activities, when it is the case both that cognitive enhancement would negatively affect the *point* of those activities (i.e. learning *through a certain kind of effort*) and that we have good reasons to value that point. However, cognitive enhancement should be *allowed* in two groups of cases, namely in practice-oriented activities (e.g. non-competitive Sudoku), on which little of moral value or social import hangs, and in some prominently goal-directed activities, such as high-responsibility professions, the goal of which has significant moral or social value. According to our approach, there are also special circumstances in which the use of efficacious and relatively safe cognitive enhancers may even be *obligatory*, typically emergency situation involving high-responsibility professionals.

We think that our reasoning may have an interest that goes beyond the ethics of cognitive enhancement, as the arguments that we put forward may also be applied - maybe with some modification - to other kinds of human enhancement. We thus hold that the structure and methodology we follow may constitute a template for a fruitful ethical discussion in this and in related areas.

REFERENCES

Gawande, A. (2012). Two Hundred Years of Surgery. *New England Journal of Medicine*, 366(18), 1716–1723.

Goold, I. & Maslen, H. (2014). Must the Surgeon Take the Pill? Negligence Duty in the Context of Cognitive Enhancement. *The Modern Law Review*, 77(1), 60–86.

MacIntyre, A. (1985). *After Virtue*. 2nd ed. London: Duckworth.

McCabe, S.E., Knight, J., Teter, C.J. & Wechsler, H. (2005). Non-medical Use of Prescription Stimulants Among US College Students: Prevalence and Correlates from a National Survey. *Addiction*, 100(1), 96–106.

McCabe, S.E., Teter, C.J. & Boyd, C.J. (2004). The Use, Misuse and Diversion of Prescription Stimulants Among Middle and High School Students. *Substance Use & Misuse*, 39(7), 1095–1116.

Murray, T.H. (2008). Sports Enhancement. In Crowley, M. (ed.). *From Birth to Death and Bench to Clinic: The Hastings Center Bioethics Briefing Book for Journalists, Policymakers, and Campaigns*. New York: The Hastings Center, 153–158.

President's Council on Bioethics. (2003). *Beyond Therapy*. https://bioethicsarchive.-georgetown.edu/pcbe/reports/beyondtherapy/beyond_therapy_final_webcorrected.pdf (Accessed March 1, 2014)

Rawls, J. (1955). Two Concepts of Rules. *The Philosophical Review*, 64(1), 3–32.

Repantis, D., Schlattmann, P., Laisney, O. & Heuser, I. (2010). Modafinil and Methylphenidate for Neuroenhancement in Healthy Individuals: A Systematic Review. *Pharmacological Research*, 62(3), 187–206.

Santoni de Sio, F., Faulmüller, N., Savulescu, J. & Vincent, N.A. (forthcoming). Why Less Praise For Enhanced Performance? Moving Beyond Responsibility-shifting, Authenticity, and Cheating to a Nature of Activities Approach. In Jotterand, F. & Dubljevic, V. (eds.), *Cognitive Enhancement: Ethical and Policy Implications in International Perspectives*. Oxford: Oxford University Press.

Savulescu, J., Foddy, B. & Clayton, M. (2004). Why We Should Allow Performance Enhancing Drugs in Sport. *British Journal of Sports Medicine*, 38(6), 666–670.

Savulescu, J. (2006). Conscientious Objection in Medicine. *BMJ : British Medical Journal*, 332 (7536), 294–297.

Schermer, M. (2008). On the Argument That Enhancement Is 'cheating'. *Journal of Medical Ethics*, 34(2), 85–88.

Searle, J. (1995). *The Construction of Social Reality*. New York: Free Press.

Sugden, C., Housden, C., Aggarwal, R., Sahakian, B. & Darzi, A. (2012). Effect of Pharmacological Enhancement on the Cognitive and Clinical Psychomotor Performance of Sleep-deprived Doctors: a Randomized Controlled Trial. *Annals of Surgery*, 25 (2), 222–227.

Vincent, N.A. (2011). The Challenges Posed to Private Law by Emerging Cognitive Enhancement Technologies. *The Law of the Future and the Future of Law*, 511–521.

Whitehouse, P., Juengst, E., Mehlman, M. & Murray, T. (1997). Enhancing Cognition in the Intellectually Intact. *Hastings Center Report*, 27(3), 14–22.

Williams, B. A. O. (1973). *A Critique of Utilitarianism*. In Smart, J. J. C. & Williams, B. A. O. (eds.), *Utilitarianism - For and Against.* Cambridge: Cambridge University Press.

Weyandt, L., Marraccini, M., Gudmundsdottir, B.G., Zavras, B.M., Turcotte, K., Munro, B. & Amoroso, A. (2013). Misuse of Prescription Stimulants among College Students: A Review of the Literature and Implications for Morphological and Cognitive Effects on Brain Functioning. *Experimental and Clinical Psychopharmacology*, 21(5), 385–407.

Is there a "Moral Obligation to Create Children with the Best Chance of the Best Life"?

Stefan Lorenz Sorgner [†]
stefan@sorgner.de

ABSTRACT

In this article, I critically deal with Savulescu's suggestion that human beings have a "moral obligation to create children with the best chance of the best life" (Savulescu & Kahane, The Moral Obligation to Create Children with the Best Chance of the Best Life, *Bioethics*, 23 (5), p. 274). I progress as follows. In part one, I will briefly describe the procedures with which Savulescu is concerned, and I will present Savulescu's argument in favour of the principle of procreative beneficence which is the basis of his argumentation in favour of the aforementioned moral obligation. In part two, I will show that the principle is inconsistent and that it violently attacks human beings who disagree with it which is the reason why I regard it as an immoral principle. In the conclusion, I will put forward some reasons for regarding the principle of procreative autonomy as morally more plausible than Savulescu's principle of procreative beneficence concerning the questions he deals with.

Introduction

Julian Savulescu claims that human beings have a "moral obligation to create children with the best chance of the best life" (Savulescu & Kahane, 2009, p. 274), because he regards the principle of procreative beneficence, abbreviated as PB, as morally right. According to this principle "couples who decide to have a child have significant moral reason to select the child who, given his or her genetic endowment, can be expected to enjoy the most well-being" (Savulescu & Kahane, 2009, p. 274). PB has been criticized by several

---

[†] Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany.

scholars during the previous 10 years (Birch, 2005; De Melo-Martin, 2004; Herissone-Kelly, 2006; Parker, 2007; Sparrow, 2007; Sparrow, 2011). In most cases the criticisms did not consider adequately his position, or implied counterarguments which are irrelevant for his line of thought. However, I think the scholars were right in rejecting Savulescu's principle of procreative beneficence. From my perspective, it is morally adequate, if one deals with challenges related to the creation of children, to take the principle of procreative autonomy into account, which is a principle Savulescu rejects.

In part one, I will briefly describe the procedures with which Savulescu is concerned, and I will present Savulescu's argument in favour of the principle of procreative beneficence. In part two, I will show that the principle is inconsistent and that it violently attacks human beings who disagree with it which is the reason why l regard it as an immoral principle. In the conclusion, I will put forward some reasons for regarding the principle of procreative autonomy, short as PA, as morally more plausible than PB concerning the questions Savulescu deals with.

1. Creating Children and the Principle of Procreative Beneficence

In the above mentioned phrase, Savulescu talks about creating children. Creating children goes beyond merely bringing new children into existence, but implies to actually do something to influence the genetic makeup of one's offspring. This can be done in various ways. Two methods are particularly prominent when technologies of genetic enhancement are being discussed: l. Creating a child by selecting a fertilized egg after in vitro fertilization and preimplantation genetic diagnosis (PGD) (genetic enhancement by selection); 2. Creating a child by actively altering a gene of an already given genetic makeup which could get done by means of transduction whereby you enter a modified virus into the cell which gradually changes a certain gene in all of someone's somatic cells (genetic enhancement by modification).

When Savulescu talks about the principle of procreative beneficence, he merely has the first option in mind, as he is claiming the following: When we have the reliable and safe option of choosing fertilized eggs after an in vitro fertilization and PGD, then we have the moral duty to choose the entity with the best chances of the best life. (Savulescu & Kahane, 2009, p. 274)

Here it can get asked whether the probability or the quality of life ought to be considered most, but this is not a crucial worry of mine. In the 2009 version

of the principle of procreative beneficence, which he formulated together with Guy Kahane, it is clear that his principle applies only to the method of selecting a child:

> If couples (or single reproducers) have decided to have a child, and selection is possible, then they have a significant moral reason to select the child, or the possible children they could have, whose life can be expected, in the light of the relevant available Information, to go best or at least not worse than any of the others. (Savulescu & Kahane, 2009, p. 274)

Savulescu does not claim that we have the moral obligation to choose the best child among the 50 or 60 fertilized eggs which were created during one in vitro fertilization process, but he has in mind that we have duty to choose the best child among the totality of children which can come about during all potential processes of in vitro fertilization. If no suitable entity is there this time, then we ought to try it one or several more times (Savulescu, 2001, p. 417).

Given his utilitarian approach (non in its act utilitarian version), he calculates as follows to get to know what is morally appropriate. He compares two events or situations, namely the overall utility before a family has created a child and after it has taken place. According to his moral philosophy, the act is morally right which brings about the most overall utility. Given his reflections, this is the case when a family has a child with the best chance of the best life, because then the overall utility will have been maximized. The example which he presents to support his line of reasoning is the following:

> Imagine now you are invited to play the Wheel of Fortune. A giant wheel exists with marks on it from $0\pm\$1000000$, in $100 increments. The wheel is spun in a secret room. It stops randomly on an amount. That amount is put into Box A. The wheel is spun again. The amount which comes up is put into Box B. You can choose Box A or B. You are also told that, in addition to the sum already put in the boxes, if you choose B, a dice will be thrown and you will lose $100 if it comes up 6. Which box should you choose? The rational answer is box A. Choosing genes for non-disease states is like playing the Wheel of Fortune. (Savulescu, 2001, p. 414)

Savulescu's example works as a thought example. If you wish to maximise your money, it is rational to choose box A. He implies that there is an analogy between his boxes and the genetic makeup's of the aforementioned case, and he assumes that in any given situation we have one state of genetic makeup which clearly has the best overall utility.

2. Counterarguments against the Principle of Procreative Beneficence

In the following section, I will initially present some general counterarguments and then some more specific and crucial ones to show that his principle is inconsistent and immoral, as it implies a cruelty and violence against minority groups which do not agree with his ethical theory.

2.1. General Worries

Let me return to his analogy between boxes and disease states. If you wish to make money, choosing Box A is the rational decision. His apparently plausible analogy between boxes and genetic makeup's, however, does not work because when we deal with genes and disease states the question of what is best is not answered that easily. Let us assume that an in vitro fertilization has taken place and afterwards some fertilized eggs underwent PGD. Consequently, we know that the fertilized egg A has characteristics a1, a2 and a3, entity B has b1, b2 and b3, and entity C has c1, c2, and c3.

We know that A will have an above average intelligence, have an excellent memory and probably be homosexual, entity B will be physiologically strong, promises to be exceptionally healthy, and will have an average height, and entity C will be exceptionally intelligent, extremely aggressive, and has a long life expectancy.

Which entity is supposed to be the best one? Is it obvious to all of us what counts as a negative trait? In addition, things are getting even more complicated, because Savulescu stresses that the principle does not only apply to several choices at one time, but potential future children also need to be taken into consideration which becomes clear when he talks about Parfit's rubella example and the case of the nuclear accident (Savulescu, 2001, p. 417-418).

Let us assume that a family wishes to have one child. Given Savulescu's PB they have the moral obligation to choose the child with the best chance of the best life. Hence, parents also need to consider the following. At time t1, they can chose between the above mentioned entities A, B and C. None of them is exceptionally intelligent, has an outstanding memory, possesses a strong health and has a very long life expectancy which nonetheless would be the desired combination of the couple in question. Consequently, they also wish to consider entities D, E, and F at time t2, which is a couple of month later, and check whether their genetic makeups are more promising. However, it is

impossible to compare entities A, B, and C with entities D, E, and F because the qualities of D, E, and F cannot be known. The option of making a comparison could only be given if we freeze fertilized egg cells A, B, and C, and compare them to D, E, and F once they will be available for making a comparison.

As you can never know whether A, B, or C is better than a later D, E, and F, parents will always be obliged to undergo a new in vitro fertilization given PB because it can always be the case that the later entities will be better than the ones currently available. When is it the case that a couple is morally allowed to stop searching for a better fertilized egg according to PB? A moral theory which demands to choose the best child in such circumstances is not a helpful theory, because it does not have a serious practicable applicability.

The principle of procreative beneficence, in principle, does not apply only to in vitro fertilizations. Understood in a wider sense, without it being restricted only to selection procedures, it is open to various processes of fertilization. The principle can also imply that couples are morally obliged to use the method of an in vitro fertilization instead of relying on the "natural" method for reproduction. In the case of sexual intercourse, the genetic makeup comes about by chance whereas in the case of an in vitro fertilization, parents can choose the best child. As the probability that the best child comes into existence is highest, when parents have the option of choosing their child, the principle implies both that we have a moral duty to avoid sexual intercourse as a method for procreation as well as a duty to use a method of contraception to make sure that sexual intercourse will not be one's own method of reproduction. However, I do not think that this is a crucial counterargument against Savulescu's principle, because he stresses that his principle is a *pro tanto* obligation which implies that it can be overruled by other insights, e.g. maybe the moral value of having a child by means of sexual intercourse.

2.2. Inconsistencies

In this section l will present some reasons for holding that Savulescu's web of thoughts associated with his PS is inconsistent. This is the case, because he refers to at least two incompatible standards of goodness within his argument. When he describes his principle of procreative beneficence he refers to specific qualities which are supposed to be associated with a good life. Thereby, he mentions being healthy, strong, intelligent, long living and having

a strong memory. However, when he replies to counterarguments against his principle he alters his concept of goodness and suddenly allows and focuses on other factors such as external circumstances and social settings which becomes particularly clear in the latest reformulation of his principle which he published in the article "The Moral Obligation to Create Children with the Best Chance of the Best Life" which he wrote together with Guy Kahane and which came out in 2009 in the Journal "Bioethics".

In one sentence, Savulescu claims that PB is neutral to central philosophical issues concerning the good life. He explicitly says that "PB is neutral with respect to such philosophical disputes about the nature of the good life." (Savulescu & Kahane, 2009, p. 279) In the paragraph before this sentence he talks about various philosophical theories of the good life. However, in the next sentence he affirms the following:

> But although there is this philosophical disagreement, there is considerable consensus about the particular traits or states that make life better or worse, a consensus that would rule out many procreative choices as grossly unreasonable [...] PB doesn't rely on some special and controversial conception of well-being. All it asks us is to apply in our procreative decisions are the same concepts we already employ in everyday situations. (Savulescu & Kahane, 2009, p. 279)

In this phrase, he makes clear that PB is not neutral to philosophical theories of the good life, but that he upholds a common sense type of approach to the good life. He is also confident to know that "there are plenty of cases where we can rank the goodness of lives. We do so in numerous moral decisions in everyday life" (Savulescu & Kahane, 2009, p. 279). According to Savulescu, we know what a good life is and what we regard as a good life actually is what is needed for a good life. Yet, he does not specify further who is the 'we', to which he is referring? Does he refer to the majority of people in Western countries, fellow intellectuals at the University of Oxford, or all strong interest groups in Western countries, like US country folk?[1]

---

[1] I would not dare to claim to know what a food life is, and I think that there are good reasons for doubting that any non-formal account of goodness is bound to be highly implausible, because any account of the good is closely connected to personal physiopsychological wishes, drives and desires. This is the reason why I regard the fights in favors of the norm of negative freedom which have taken place during the Enlightenment as praiseworthy events (Sorgner, 2010, p. 240).

In a further paragraph, he slightly alters his concept of the good life again, because he puts forward that it is supposed to be clear that there is such a thing, not only as the good life, but even as the best life, as he claims:

> A common objection to the PB is that there is no such thing as a better or best life. It is hard to defend such a claim. (Savulescu & Kahane, 2009, p. 278)

The validity of there being a good life can imply that there are some general principles which are valid for all human beings, which is already an extremely strong claim. However, Savulescu in this phrase moves beyond the affirmation of the good life towards the belief that there is actually a best life. Plato has held such a position which he manages to explain by reference to his theory of forms. However, for a libertarian, utilitarian philosopher such as Savulescu to uphold this claim is quite a daring position. I would be keen to know how he can manage to explain this theoretical foundation of the PB.

Savulescu does not only make general remarks concerning the question of the good, but actually puts forward a list of qualities what we regard as good and bad, according to his perspective: It is bad to have a disposition for depression (Savulescu & Kahane, 2009, p. 281) and having a disability (Savulescu & Kahane, 2009, p. 286). It is good, on the other hand, to have good memory, and a strong intelligence (Savulescu, 2001, p. 420), according to in his paper from 2001. In his paper from 2009 having a strong memory and being able to concentrate well and understand other people's feelings seems central for him, as he holds:

> How can the capacity to remember things better, concentrate longer, be less depressed, or better understand other people's feelings have the effect that one will be less likely to achieve the good life? (Savulescu & Kahane, 2009, p. 284)

Having a high intelligence is also helpful for a good life, according to him:

> If parents could increase the prospects off future children's lives by selecting children who are far more intelligent, emphatic or healthier than existing people, than PB instructs parents to select such future children. (Savulescu & Kahane, 2009, p. 290)

I am not entirely certain concerning all of the implications of his theory of the good. Does he mean that a human entity is always better off with a higher intelligence, a stronger memory or more intensive capacity to concentrate? I am doubtful whether this is actually the case. It is good to have good memory and I definitely wish to have a stronger memory. On the other hand, I am also

grateful for having a good capacity to forget things. If I permanently remembered all the bad things friends have done to me or were conscious of dangers related to what I am doing, then I would most probably to decide to move away from civilization and live like a hermit on the top of a mountain. Hence, I would be hesitant to claim that having a better memory necessarily is always a good thing, and that we have a moral reason for always choosing the fertilized egg with the better memory. Likewise, it can get asked whether it is always in the interest of a child to have a higher intelligence. If the child is the only one with such a high intelligence, while all the people around him are fools, then I would be hesitant to claim that having such a high intelligence necessarily is in the best interest of the child. Hence, it seems to me that the social settings are of central relevance for the qualities necessary for leading a good life.

Savulescu seems to have realized and acknowledged the impact of this line of thought in between 2001 and 2009 especially with respect to the question of disability which can and ought to get transferred also to other domains of a good life. In his article from 2001 Savulescu upholds the following position:

> The reason is that it is bad that blind and deaf children are born when sighted and hearing children could have been born in their place. (Savulescu, 2001, p. 423)

In his 2009 paper he developed his views concerning disabilities further:

> In this final section we shall argue that PB provides a better approach to the question of disability than the competing procreative principles. (Savulescu & Kahane, 2009, p. 284)

According to Savulescu and Kahane, it needs to be stressed that "disability is a context and person-relative concept. What may make it harder to lead a good life in one circumstance may make it easier in another" (Savulescu & Kahane, 2009, p. 286).

As a consequence of his altered approach concerning disability, they stress that:

> on our account of disability, people do have reasons not to have a future child who is likely to be disabled if they have the option of choosing another who is expected to less of no disability, although whether it would be wrong to do so would depend on the overall balance of moral reasons. (Savulescu & Kahane, 2009, p. 286)

They even reach the following conclusion given their new approach:

> If a case can be made that deafness is not a disability in our sense— if it can be shown that deafness does not reduce well-being, or at least that in a given context deafness is not expected to be a disability, then PB would not give any moral reason not to select deafness. (Savulescu & Kahane, 2009, p. 289)

Their final remarks concerning disability are actually quite interesting and do have some plausibility. However, I wonder which of his remarks represent his theory of the good life upon which his PB is based. In his writings, it is possible to find a wide range of affirmative statements concerning the good life which are mutually incompatible. It seems as if he has got a theory of the good for every difficult question from his critics. On the one hand, he upholds a perfectionist theory of the good life which can be identified with being more intelligent, healthier, having a stronger memory and so on. On the other hand, he refers to what we regard as good life, which can be referred to as a common sense approach to the good life, and I am pretty certain that it would not be a perfectionist theory of the good, which would get upheld, if large groups of Western citizens were asked what a good life is. My assumption gets support from the fact that many US mothers who ordered sperm were interested in sperm from good looking, and sportive Ivy League students rather than in the sperm of Noble price winners. (Caplan, 2012, p. 156) Finally, his statement that PB is neutral to classical philosophical theories of the good life has to be mentioned, too.

I am not able to order his various utterances such that they fit together consistently. It rather seems to me the case that he upholds various mutually exclusive theories of the good which he uses in order to have plausible replies to the worries of his critics.

2.3. PB as a Violent and hence an Immoral Principle

According my point of view, PB is not only inconsistent, but it is also an immoral principle because it acts violently against individuals and interest groups who do not agree with PB and the associated theory of the good or should I say the corresponding theories of the good. Savulescu himself is aware that PB is a much stronger theory than the theories of the good which most liberal ethicists have proposed in recent years, and he regards it as morally appropriate that his is such, as he clearly holds the following view:

> Although PB and the procreative principles we have considered here bear little resemblance to the collectivist, coercive, and often racist projects of 20th

> Century eugenics, most supporters of genetic selection have tended to proceed
> gingerly, defending views that are unnecessarily weak. (Savulescu & Kahane,
> 2009, p. 282)

I think that Savulescu is right in stressing the difference between his PB and
procedures of state governed eugenics during the Third Reich, even though
some commentators have criticized him for proposing a new type of eugenics
(e.g. Sparrow, 2011). However, you do not have to propose Fascist views for
holding an immoral view. He even regards PB as a moral principle:

> PB is a moral principle. It states what would be morally right or wrong for
> reproducers to do. (Savulescu & Kahane, 2009, p. 279).

Let me put forward some reasons for regarding PB as immoral. Firstly, I regard
it as immoral because it is violent against people who do not subscribe to it.
Secondly I see it as immoral because it implicitly contains immoral duties.

In how far can it be said that PB acts violently against people who do not
subscribe to it? PB acts violently, because it implies that parents act immorally
who do not subscribe to the theory of the good life as Savulescu proposes it.
The principle demands to tell these parents that they ought not to have acted
the way they did, and that thereby they have acted falsely. In this way, the
principle intrudes paternalistically in the life of other people and acts violently
against their concept of the good life.

Parents might decide not to choose the fertilized egg which has the greatest
memory, because they know that the child will grow up in a poor family and
during war times and they think that it will most probably be good for the child,
if it does not have to remember all the bad things which are bound to take place
given such problematic circumstances.

PB implies that if parents prefer someone with a weaker memory to
someone with a stronger memory, they act immorally. I do not think that the
parents acted immorally in these circumstances, but rather that they have made
a decision concerning the good life which is understandable given the
circumstances they live in.

Savulescu might reply that PB allows parents to consider the circumstances
in the process of evaluating which qualities increase the child's probability of
living a good life, as he dealt with this issue analogously in the above case of
disability. Still, above he also made clear that this position is valid only for the
case of disabilities and given the other phrases cited before he holds that having
a stronger memory is better than having a weaker one. If in the given context he

claims that the social circumstances ought to be considered, too, it becomes clear that he holds mutually exclusive concepts of the good life. If he claims that the choice of the parents in question was immoral, he upholds an immoral position, because he acts violently against the view of the good of the parents and acting violently in this manner is an immoral act.

Besides the PB being violent, it also implies immoral duties. Why is this so? It is the case, because Savulescu fails to see the impact of a distinction to which he himself referred in his article on procreative beneficence from 2001, namely the distinction between genetic enhancement by means of selection and by modification, e.g. by modifying an already given genetic makeup. The first type of procedure is structurally analogous to choosing a partner with whom one wishes to have offspring. The second type of procedure, however, is structurally analogous to educating ones offspring. (Sorgner, 2013, p. 85-100).

In the following section I will merely state a few reasons in favour of my claim that selecting an already given genetic makeup and selecting a partner with whom one wishes to have offspring are structurally analogous procedures (see Sorgner, 2011, p. 21-25).

By choosing a partner with whom one wishes to have offspring, one thereby implicitly also determines the genetic makeup of ones kids, as 50 per cent of their genes come from ones partner, and the other 50 per cent from oneself. By selecting a fertilized egg, one also determines 100 per cent of the genetic makeup by means of selection.

One objection, which might be raised here, is that selecting a fertilized egg cell is a conscious procedure but normally one does not choose a partner according to their genetic makeup such that one has specific genes for one's child. However, it can get replied that our evolutionary heritage might be more effective during the selection procedure of a partner than we consciously wish to acknowledge. In addition, the qualities according to which we choose a fertilized egg after a PGD might not have been chosen as consciously as we wish to believe, but might be influenced more on the basis of our unconscious organic setup than we wish to acknowledge. It might even be the case, that the standards for choosing a partner and for choosing a fertilized egg might both be strongly influenced by our organic makeup and evolutionary heritage such that both are extremely similar.

A difference between these two selection procedures is surely that in the one case, one selects a specific entity, a fertilized egg, but in the other case a partner and therefore only a certain range of genetic possibilities. However,

given the latest epigenetic research, we know that genes can get switched on and off which makes an enormous difference on the phenomenological level. Hence, it is also the case that by choosing a fertilized egg, we only choose a certain range of phenomenological possibilities of the later adult, as is the case by choosing a partner for procreative purposes.

The aforementioned comparison provides some initial evidence for holding that there is a structural analogy between choosing a partner for procreative purposes and for choosing a fertilized egg cell after PGD. Given PB and given that structurally analogous procedures ought to be evaluated analogously, PB implies not only that there is a moral duty to select the child with the best chance of the best life, but also that there is the moral duty to select the partner for having offspring such that the likelihood is maximized that a child with the best chance of the best life can get realized. Anyone who does not stick to this moral duty of his can get told off and be told that he ought not to behave in this way and that he is acting immorally.

A moral principle, and PB claims to be such a principle, which implies the moral duty to select a partner with whom ones offspring promises to be best is extremely violent, and hence immoral.

3. Conclusion

The main goal of my text was to deal with the question whether we have a "Moral Obligation to Create Children with the Best Chance of the Best Life?". After having dealt critically with Savulescu's PB which claims that there is such an obligation, I conclude that his arguments in favour of such a duty fail, as they are inconsistent and immoral. From this it does not yet follow that there is no such duty, but it merely means that his arguments in favour of such a duty are implausible. Without having the time and space to move beyond this conclusion, I wish to point out that I regard the principle of procreative autonomy as an appropriate one. Savulescu has argued in various articles that PA is not the appropriate attitude with respect to the process of selection after in vitro fertilization and PGD. According to Savulescu, procreative autonomy can be summarized thus:

> Procreative autonomy. If reproducers have decided to have a child, and selection is possible, then any procreative option selected by reproducers is morally permissible as long as it is chosen autonomously. (Savulescu & Kahane, 2009, p. 279)

Given that that there is a structural analogy between choosing a partner for procreative purposes and choosing a fertilized-egg cell after PGD, I regard procreative autonomy as morally appropriate for a liberal state. Here, the individual's right to live a good live according to his own concept of a good life is of central importance. During the previous 500 years, enormous and intensive fights in various social and political fields have taken place until negative freedom has been widely recognized as a central norm, and I regard it as important to always take this achievement into consideration. A move away from procreative autonomy towards procreative beneficence is a move into the wrong direction, because it introduces new paternalistic structures. This time the structures are not given on a legal level, but merely on a moral one. Still, it has the effect that such structures violently intrude in the private realm of individuals and violently attack the precious achievement that it is widely recognised that a radical multiplicity of concepts of the good can be appropriate. Hence, I finally conclude that not only do we not have a "Moral Obligation to Create Children with the Best Chance of the Best Life", but I am even bound to claim that it is immoral to defend a "Moral Obligation to Create Children with the Best Chance of the Best Life" or in other word: PB is not a moral but it is an immoral principle.

REFERENCES

Birch, K. (2005). Beneficence, Determinism and Justice: an engagement with the argument for the genetic selection of intelligence. *Bioethics*, *19*(1), 12–28.

Caplan, R. (2012). *Selfish Reasons to Have More Kids: Why Being a* Great Parent is Less Work and More Fun Than You Think. New York: Basic Books.

De Melo-Martin, I. (2004). On our obligation to select the best children: a reply to Savulescu. *Bioethics*, *18*(1), 72–83.

Herissone-Kelly, P. (2006). Procreative Beneficence and the prospective parent. *Journal of Medical Ethics*, *32*, 166–169.

Parker, M. (2007). The Best Possible Child. *Journal of Medical Ethics*, *33*(5), 279–283

Savulescu, J. (2001). Procreative Beneficence: Why we should select the best children. *Bioethics, 15*(5–6), 413–426.

Savulescu, J. and Kahane, G. (2009). The Moral Obligation to Create Children with the Best Chance of the Best Life. *Bioethics*, *23*(5), 274–290.

Sorgner, S. (2010). *Menschenwürde nach Nietzsche. Die Geschichte eines Begriffs*. Darmstadt: WBG.

Sorgner, S. (2011). Zarathustra 2.0 and Beyond. Further Remarks on the Complex Relationship between Nietzsche and Transhumanism. *The Agonist*, Vol. IV, Iss. II, Fall 2011, 1–46.

Sorgner, S. (2013). Evolution, Education, and Genetic Enhancement. In: Sorgner, Stefan Lorenz and Jovanovic, Branka-Rista (eds.): *Evolution and the Future: Anthropology, Ethics, Religion*. Frankfurt a. M. et al.: Peter Lang, 85–100.

Sparrow, R. (2007). Procreative Beneficence, Obligation, and Eugenics. *Genomics, Society and Policy* 2007, Vol.3, No.3, 43–59.

Sparrow, R. (2011). A not so new eugenics: Harris and Savulescu on human enhancement. *Hastings Center Report*, *41*(1), 32–42.

Economic Inequality and Human Enhancement Technology

Fox Swindells[†]
fox.swindells@pg.canterbury.ac.nz

ABSTRACT

Human Enhancement Technology ranges from the commonplace, such as education, to the futuristic, with possible future developments including genetic modification or direct computer-brain interfaces. Public policies governing the supply of these technologies have the potential to greatly increase or mitigate economic inequality. Due to this potential harm, many have suggested prohibition of further developments of enhancement technologies. However, prohibition would in ineffective at preventing this harm and also would also prevent many positive aspects of enhancement technologies. On the other hand, due to the expected benefits, many have suggested allowing access and development within a free-market system. However, this has the potential to increase inequality beyond acceptable levels. Consequently, Government policies must provide appropriate funding and regulation in order for these technologies to be distributed fairly to provide the most benefits and prevent the worst outcomes.

Introduction

Human Enhancement Technology has the potential to provide both great benefits and greater inequalities if left unchecked. Because of this, we must consider a variety of regulatory policies in order to achieve the best outcome. The most common options to be considered include prohibition on the technology and its developments, allowing access within a free-market system, or government distribution. Prohibition would probably be ineffective and undesirable, and a free-market system would likely result in the greatest

[†] University of Canterbury at Christchurch, New Zealand.

inequalities and negative outcomes for all involved. Therefore, a compromise is the best alternative; access to human enhancement should be ensured for all in order for the best outcome to be achieved. Government funding and regulation can ensure low cost and equal distribution of human enhancement technologies, and consequently should be implemented as the best policy option.

1. The Technology

Human Enhancement Technology (HET) encompasses a wide range of technologies, from the commonplace to the futuristic. Although the most powerful enhancements may not yet exist many are already available. For example, education and caffeine are not commonly considered HET, but they enhance individuals' capabilities and, therefore, can legitimately be considered HET and directly compared to new HET. Education teaches mental software for managing cognitive domains to reduce mental load (Sandberg & Bostrom, 2007, p. 208), and caffeine is used by millions of adults daily for its stimulant effect (Bramstedt, 2007, p. 1237). There is also a long history of using external hardware to increase cognition, such as pen and paper or personal organizers, and this use is constantly increasing with smart phones, virtual reality, and direct computer-brain interfaces. Many current, generally low-tech, HETs are well accepted, for example, most people have no problem with individuals using caffeine or education and they do not consider it an unfair advantage as it is commonplace (Sahakian & Morein-Zamir, 2007, p. 1158). These technologies are motivated by the possibility of enhancing human capacities beyond what the average human is naturally capable of. Consequently, enhancements are constantly increasing in their ability to improve capabilities. Future HETs have the potential to allow the brain to learn quickly and improve selective retention, unlearn phobias and addictions, increase fine-grained control over the learning process, increase creativity, and improve memory (Sandberg & Bostrom, 2007, p. 203–207). Francis Fukuyama (2002, p. 8–9) believes that HET will allow us to change our personality, grow new organs, repair our brains, and extend life expectancy beyond 100 years.

Although increasing human capabilities is the goal of enhancement, many HETs were not originally developed for this purpose. Many were developed as therapies for disabilities. For example Ritalin, which was developed as a

treatment for attention deficit hyperactivity disorder, is used by college students to enhance their cognition (Lamkin, 2012, p. 347), and Modafinil, originally developed as a treatment for narcolepsy, is used to reduce performance decrements from sleep loss or jet lag (Sandberg & Bostrom, 2007, p. 204). Consequently, the exact HETs that will be developed are unclear, and this means predicting the outcomes of their development is difficult. However, one outcome is likely to apply to any enhancements developed; they will be expensive and, therefore, only affordable to the better-off in society if left unregulated. Based on this assessment, I will consider all HET together to assess this consequence, as a response combating this harmful outcome must be created prior to their development.

2. Increased Inequality

The expected inequality of access to HET will exacerbate existing economic inequalities if left unregulated. The wealthy already benefit from their financial situation; for example they can use their position to access better education and nutrition, which in turn enhances their brain power (Sahakian & Morein-Zamir, 2007, p. 1159). HET has the potential to allow those who can afford it to increase, through the use of genetic modification technology, their own, and their children's, IQ beyond even that of the most gifted naturally. Cognitive enhancements, such as education, have many benefits beyond higher job status and salary; they can reduce the risk of substance abuse, crime, and many illnesses while increasing quality of life, social connectedness, and political participation (Sandberg & Bostrom, 2007, p. 208). Consequently, the benefits associated with higher IQ, such as increased income (Sandberg & Bostrom, 2007, p. 216), and prevention of a wide array of social and economic misfortunes (Bostrom & Sandberg, 2009, p. 330), are likely to increasingly become solely available to those who are better-off, further increasing the advantages packaged with wealth. This will exacerbate economic inequality by providing further benefits to those with the ability to pay and preventing access for the less well-off.

Fukuyama (2002, p. 9-10) is concerned that the idea of natural human equality, that is the base of political and moral equality, will be compromised by HET and consequently some people, the unenhanced, will be considered less human than the enhanced. HET has the potential to create two classes of people, the enhanced and the unenhanced, and this would increase class

struggle; a solid immovable hierarchy would form where, based on ability to pay, some people would be significantly better off than others who would never have the ability to catch up as they lack class mobility. There is concern that those able to afford HET will be buying their own well-being at the expense of a greater social good (Caplan & Elliott, 2004, p. 174). There is fear that the way we live together as a group could be damaged by the actions of individuals. Harms from inequality do not require extreme deprivations to warrant our consideration; injustices exist even when no extreme deprivation is present. If a HET increased political influence for those who could afford it, such as by allowing increased communication capacities, this would be an injustice to those who did not have access to it although they suffer no extreme deprivation (Buchanan, 2011b, p. 250). We must seriously consider these potential harms from increased inequality and create policies to best mitigate these harms.

Studies have found a wide range of negative outcomes both within and between nations with greater inequality, these include; greater risk of mental disability and psychiatric hospitalization (Hudson, 2005, p. 16); lower economic mobility (Andrews & Leigh, 2009, p. 1492); poorer general health; higher infant mortality; lower average life expectancy; increased obesity; greater illicit drug use; higher homicide and violent crime; a greater prevalence of depression; and, lower self-reported well-being (De Vries, Gosling & Potter, 2011, p. 1978). These numerous social problems are more common in unequal societies, for everyone in the society, not just the less well-off (Wilkinson & Pickett, 2007, p, 1972). In a society with a strong hierarchy, an individual's position relative to others is more important, and, consequently, individuals become more competitive, less trusting, more self-focused, less friendly, and less cooperative (De Vries, Gosling, & Potter, 2011, p. 1979). This means more unequal societies have lower levels of agreeableness and, following from this, poor health outcomes, such as poor diet, and increased alcohol and cigarette consumption (De Vries, Gosling, & Potter, 2011, p. 1984). Increased inequality, and the associated negative consequences, should be of concern to both the less and more well-off in society.

3. Prohibition

Because of these potential harms some might suggest that we should prohibit HET in order to avoid the inescapable inequality that seems to be bundled with its development. However, prohibiting or severely restricting HET, or at least

any future development, is likely to be ineffective. Prohibition is likely to be an ineffective method for controlling the outcomes. Any prohibition is likely to push HET underground and across boarders. The probable outcome from this is price increases that further widen the gap between those who can afford HET and those who cannot (Stock, 2005, p. 29). Similarly, even if prohibition were successful in some countries, other areas lacking the prohibition will continue with the developments. Even if those developments were prevented from going to the countries with prohibition the areas without the prohibition would have an advantage because of their access to HET. Pharmaceuticals will also continue to be developed as therapeutic drugs in many locations, even those with prohibition on HET, but these often also have enhancement abilities, and preventing 'off-label' uses is impractical (Buchanan, 2011a, p. 158). Therefore, attempting to implement prohibition of HET would be a seemingly futile effort.

Although prohibition may be ineffective, this is not sufficient reason to not introduce one if it is the best option. For example, we have laws against murder even though it is sometimes ineffective as a method of prevention. Effectiveness, or its lack, is not alone a sufficient reason to support or oppose a policy. Other expected outcomes of the policy must be considered. HET has the potential to eliminate, not just increase, many social problems by allowing increased control over aspects of our personalities, such as prejudice, discrimination, laziness, apathy, cruelty, anger and to also make people smarter, more insightful, and more athletic (Borenstein, 2009, p. 521). HET could reduce inequalities and provide positive benefits in many ways, not only cause harm. For example, much of human cognition is shared between minds and more efficient forms of collaboration, such as virtual workspaces and internet which are used already, can therefore enhance cognition (Sandberg & Bostrom, 2007, p. 213). Although there is little evidence that greater intelligence causes greater happiness, there is evidence that higher intelligence increases health and wealth, while lower intelligence puts an individual at greater risk of accidents, negative life events, and low income (Sandberg & Bostrom, 2007, p. 201). Increased cognitive ability helps individuals tackle the increasingly complex demands society places on cognition. Increased cognitive ability is not only a positional good, it is also intrinsically valuable and its value does not depend on other people lacking it; for example, having a good memory or increased creativity is valuable even if others have a similar level of excellence (Bostrom & Sandberg, 2009, p. 328). These cognitive

abilities are valuable for society as well as individuals; many social problems could potentially be solved if people were smarter, wiser, or more creative. Alan Buchanan (2011b, p. 247) argues that HET could promote justice because many HETs potentially work better for those of lower cognitive ability and could be cheaper than educational interventions. HET could also be used to remove disabilities and generally increase individuals' abilities in many ways (Buchanan, 2011b, p. 428), for example, HET has the ability to remove deafness or increase hearing, as cochlear implants already do today. Prohibition of HET would limit the socially beneficial uses whereas legal enhancements have the ability to lead to safer, cheaper, enhancements. HET has the potential for many outcomes, some negative and others overwhelmingly positive. If we prohibit HET it is unlikely that we will prevent the negative consequences, but it is certain that we will prevent many, if not all, of the positive outcomes.

4. Free-Market Distribution

Conversely to those arguing for prohibition of HET, it is understandable that enormous potential benefits lead some to favor allowing as much access and development as possible. A free-market system is likely to be the favored method for providing this access, it would mean that those who can afford the technology will have access, and development is based on their demand. This is similar to how many advantages are currently distributed. We allow more well-off individuals to enjoy many advantages over their less wealthy peers, with very few policies that prevent them from doing so. For example, we allow wealthy students to employ private tutors or to have more time for study because they do not have to work to support themselves. We do not hold that justice demands enhancements should not be available to any until they are available to all. If this were the case, we would require literacy campaigns to halt in countries with high levels of literacy until all countries catch up with those ahead (Buchanan, 2011a, p. 158). We already find it acceptable for some to have access to enhancements although others do not.

   If HET were distributed through a free-market system, this would obviously be similar to the current situation of many technologies, and consequently we can compare the expected consequences to those currently experienced. It can be expected that if HET was made available through the free-market system, prices would fall dramatically in the future when the enhancements come off

patent and generics became available. Currently, many generic versions of prescription drugs are available at a much lower cost than they were previously, and although they may have been out of reach of some in the beginning, the price has not remained high and the developments are still beneficial (Buchanan, 2011a, p. 158), for example Penicillin was originally prohibitively expensive and now is available to millions for only cents per dose. Although this price drop seems likely to happen for HET as well, the time frame is unpredictable and the wealthy will still be at an advantage as they will move on to the next enhancement developed, that will still be expensive, while those less well-off will only have access to the older, less effective, enhancements once their price has dropped.

Within the free-market system there are other potential ways a technology could be distributed to those unable to pay by their own means. For example, if a specific enhancement is considered necessary by employers, then it is likely that they will provide it for their employees, as they currently do with staff computers. However, even if the employers will supply enhancements for their employees, experience with the technology will increase an individual's employability for a position. With computers currently, proficiency with common programs is a requirement for many employment opportunities, and, consequently, those able to afford their own computer are more employable than those unable to do so. The cost of computers has finally fallen to a price range affordable to almost everyone, except the least well of, in the developed world. This means that, although most people have access now, this was not always the case. Some individuals still lack proficiency with computers, as they are unable to afford their own, which makes them less employable. Although the price of HET is likely to fall similarly, and make many HETs accessible to almost everyone, this dispersion could be slow or limited and, consequently, produce more injustices. Those who lack access in the time it takes for the technology to disperse may be unjustly excluded from important forms of political and economic participation and those with access to HET will gain many advantages (Buchanan, 2011b, p. 253). This inequality is an unacceptable consequence from the free-market system of dispersion. Even if the inequality created would not be permanent, as the price would eventually drop (which is unclear in itself), the consequences of delayed access to HET are sufficient to require action be taken to reduce harmful inequality.

5. Government Distribution

Finding the balance where the greatest benefits are available and the worst harms are prevented requires a compromise between prohibition and a free-market system. Depending on how public policy is approached, HET can increase inequality through pushing the technology underground, increasing prices, and only allowing the rich access, or it can reduce inequality by supporting responsible development and ensuring broad access. Without public funding and support, it is likely that HET will be out of reach of many, and the divide between those who are 'normal' and those who are above average will continue to grow and be based on ability to pay. Buchanan (2011b, p. 246-247) argues that the requirements of justice mean that socially produced goods, and their impacts, should be distributed. Although distributing HET will not remove all injustices, this does not mean that it is not a valuable goal; we can tolerate some injustices persisting without accepting others. Some we accept because we acknowledge that there is little we can do about them; for example, we accept differences in the amount of time students have to study because it would be difficult or impossible to enact regulations to constrain it (Lamkin, 2012, p. 349). Even if we were able to do something about these disparities, our motivation is most likely to be to provide more access for poor students rather than take access away from rich students. The reasons we have for supporting education as a public endeavor translate simply to other HET. Like education, other HETs increase an individual's well-being and better equip them for their role as a citizen (Buchanan, 2011a, p. 147). Subsiding HET would provide a public good and be a more constructive approach than other policy options, as it would speed up dispersion and help ensure that these valuable innovations quickly become widely available.

Some may object to government dispersion as the solution, arguing that our poor track record of helping the disadvantaged shows that it is likely that the rich will still have access while the poor will not. Similarly, they may argue that there may be so many HETs that it would be impossible for the government to fund all of them due to its limited resources. Although this objection is worth considering, it seems plausible that the main reason that we have failed to help the disadvantaged in the past is due to wishing to avoid spending money in this fashion, and, although the economic cost of a policy is important, it is likely that the cost of enforcing prohibition would be greater than that of providing access to HET (Lamkin, 2012, p. 350). Beyond this, it

is likely that the cost of provision will be outweighed by the benefits of increased economic and social advantages from providing HET to all. HET has the potential to reduce costs for the government in other areas, such as health care. It is likely to also be beneficial for our economy by increasing citizens abilities and national productivity. Public funding is also likely to drive down prices of HET as large corporations compete for government contracts to be the HET provider for the nation.

The potential distributive problems for HET are not novel, as with other innovations policies can worsen or mitigate inequalities. If HET is treated as a social good, as education is currently, it is likely that at least basic HET will be publicly distributed and subsidized, rather than solely available based on an individual's ability to pay (Buchanan, 2011a, p. 148). Not only does reducing inequality have positive health and social outcomes for all members of society; but also the benefits of technology are generally greater when more people have access. For example, cognitive enhancements have network effects, where the benefit increases as more individuals have the enhancement; to be more precise, being literate or having computer access is much less valuable if only a few people have those enhancements (Buchanan, 2011a, p.149). Public policies that increase the distribution of HET would be beneficial for all members of society, rather than just those who would otherwise lack the ability to pay, and consequently subsidizing access and regulating development to ensure equal access is the best option for everyone involved.

Conclusion

HET has the potential to provide many benefits to both individuals and society provided that it is fairly distributed. This requires public funding and regulations in order to avoid the worst inequalities. The obvious benefits from HET provide ample evidence for why HET should not be prohibited, and, rather, governments should fund access for all citizens to ensure that the benefits are distributed as equally as possible. Based on the expected benefits and harms from HET, public policies must be developed to ensure the best of all possible outcomes. Neither prohibiting HET or accepting access through a free-market system are effective or productive solutions as both these approaches will inevitably increase inequality. The best solution for controlling the consequences from HET is a compromise between no access and access only based on ability to pay, this option is best not just for the less well-off but also for the wealthy. Therefore, the government should

ensure distribution of HET through public funding, and regulations on development and patents, that ensure lower costs and equal access.

REFERENCES

Birch, K. (2005). Beneficence, Determinism and Justice: an engagement with the argument for the genetic selection of intelligence. *Bioethics*, *19*(1), 12–28.

Caplan, R. (2012). *Selfish Reasons to Have More Kids: Why Being a Great Parent is Less Work and More Fun Than You Think*. New York: Basic Books.

De Melo-Martin, I. (2004). On our obligation to select the best children: a reply to Savulescu. *Bioethics*, *18*(1), 72-83.

Herissone-Kelly, P. (2006). Procreative Beneficence and the prospective parent. *Journal of Medical Ethics*, *32*, 166-169.

Parker, M. (2007). The Best Possible Child. *Journal of Medical Ethics*, *33*(5), 279 - 283

Savulescu, J. (2001). Procreative Beneficence: Why we should select the best children. *Bioethics*, *15*(5-6), 2001, 413–426.

Savulescu, J., & Kahane, G. (2009). The Moral Obligation to Create Children with the Best Chance of the Best Life. *Bioethics*, *23*(5), 2009, 274-290.

Sorgner, S.L. (2010). *Menschenwürde nach Nietzsche. Die Geschichte eines Begriffs*. Darmstadt: WBG.

Sorgner, S. (2011). Zarathustra 2.0 and Beyond. Further Remarks on the Complex Relationship between Nietzsche and Transhumanism. *The Agonist*, Vol. IV, Iss. II, Fall 2011, 1-46.

Sorgner, S. (2013). Evolution, Education, and Genetic Enhancement. In: Sorgner, Stefan Lorenz and Jovanovic, Branka-Rista (eds.): *Evolution and the Future: Anthropology, Ethics, Religion*. Frankfurt a. M. et al.: Peter Lang, 85-100.

Sparrow, R. (2007). Procreative Beneficence, Obligation, and Eugenics. *Genomics, Society and Policy* 2007, Vol.3, No.3, 43-59.

Sparrow, R. (2011). A not so new eugenics: Harris and Savulescu on human enhancement. *Hastings Center Report*, *41*(1), 32– 42.

How Liberal is (the Liberal Critique of) a Liberal Eugenics?*

Nathan Van Camp[†]
nathan.vancamp@ua.ac.be

ABSTRACT

This article critically surveys the current bioethical and politico-philosophical debate about the ethical permissibility of a so-called 'liberal eugenics' and argues that neither the liberal argument for nor the liberal argument against human genetic enhancement is internally consistent as, ultimately, each ends up violating the very liberal principles it nonetheless pretends to defend. In particular, it will be shown that while the argument against a new eugenics necessarily entails a preemptive dehumanization of any potential enhanced form of life, the argument for it threatens to reduce any non-enhanced form of life to a "wrongful life" or a life not worth living. It will therefore be concluded that the specific stakes of this contentious issue cannot be grasped within a liberal conceptual framework.

Introduction

Recent progress in molecular biology and genetics has opened up the way for the deliberate manipulation of the human genome. Although there are still numerous technical barriers that have to be overcome before human genetic modification will become a standard medical procedure "the question is no longer whether we will manipulate embryos, but when, where, and how" (Stock, 2003, p. 2). The most direct benefit of genetic technologies will be in the prevention and healing of disease. But in addition to this obvious use, it will also be possible to employ these technologies for the purpose of human

---

[†] Department of Philosophy, University of Antwerp, Belgium.

genetic enhancement. That is, if we can identify the gene(s) for phenotypic characteristics like height, strength, intelligence, and temperament, then it will be possible to use this knowledge to design human beings according to our personal preferences.

In bioethical and politico-philosophical debates about genetic technologies, these developments are usually framed in terms of a return of eugenics. The central assumption guiding much of the literature on the subject is that if genetic technologies produce eugenic effects, then they are also morally unacceptable. Most recently, however, some commentators have taken a different approach to this issue. They argue that there is nothing intrinsically wrong with the goals of eugenics as such and that its moral acceptability depends on the values and principles of the political ideology that regulates its implementation in society (Agar, 2007; Dworkin, 2000; Harris, 2007). They reject as unjustified any comparison that might be drawn between the project of human genetic enhancement and earlier morally reproachable eugenic practices by arguing that the new eugenics will be firmly rooted in the core liberal principles of state neutrality and individual autonomy. Liberal critics of human genetic enhancement, on the other hand, claim that this attempt to integrate the eugenic ideal into a liberal framework is bound to fail and that it will inevitably corrupt the central tenets of political liberalism to the point of its becoming something different altogether (Fukuyama, 2002; Habermas, 2003; Sandel, 2008).

In this article, it will be argued, however, that both liberal responses to the challenge of human genetic enhancement are internally inconsistent, as both are bound to lead to conspicuously illiberal conclusions. More specifically, it will be shown that while the liberal argument in favor of enhancement threatens to deprive all non-enhanced forms of human existence from any intrinsic value, the liberal argument against enhancement threatens to do exactly the same with regard to all future enhanced forms of human existence

1. The Liberal Eugenics

It is notable that few advocates of the new eugenics are willing to call the practice they support by that name. John Harris, for example, prefers to speak of "deliberate selection" (Harris, 2007, p. 4) and Gregory Stock favors the term "human self-design" (Stock, 2002, p. 3). These authors' reluctance to use the term 'eugenics' obviously has much to do with the dark shadow that still

hangs over earlier attempts to make improvements to the biological foundations of human existence. Many critics are indeed afraid that the emergence of a new eugenics will also prompt the return of some of the horrific acts committed in the field's name, such as the atrocities committed by Nazi eugenicists. This association is so strong that occasionally, when a practice is referred to as eugenic, it is in fact being described as morally reproachable (Wilkinson, 2008; Paul, 1992). Apparently, the mere use of the label is enough to indicate that it refers to a field of practices that any reasonable person would find morally objectionable. Conversely, those who argue in favor of a new eugenics are almost invariably accused of offering a thinly veiled justification of Nazism.

Yet despite its inglorious history, the concept of eugenics continues to attract enthusiastic supporters. There appears to be something undeniably appealing in the essential idea of eugenics, something that prevents us from rejecting it in its entirety. Who, after all, would not want to give his or her child the best possible genetic endowment? Convinced that the potential benefits of genetic technologies in human reproduction are too valuable to renounce on the basis of past abuses, advocates of a new eugenics therefore argue that the main question is not whether the Nazi eugenics program was abhorrent but whether the atrocities committed in the name of eugenics were not in fact the result of the underlying Nazi ideology rather than something intrinsic to the field of eugenics itself. Provided that the eugenic goal of 'enhancing' human beings still enjoys universal support and approval, and that the moral acceptability of eugenics depends on the values and principles of the political ideology regulating its implementation in society, then, they suggest, it might still be possible to devise a form of eugenics that is compatible with the basic tenets of contemporary liberal democracy.

Nicholas Agar, one of today's most vocal advocates of a new eugenics, has argued that the central principles of liberalism provide ample guidance for avoiding the moral pitfalls of earlier forms of eugenics: "[T]he addition of the word 'liberal' to 'eugenics' transforms an evil doctrine into a morally acceptable one" (Agar, 1998, p. 135). In his view, the most important difference between the authoritarian eugenics of the past and the liberal eugenics he envisages is simply the degree of control that the state has over the reproductive choices of its citizens: "While old fashioned authoritarian eugenicists sought to produce citizens out of a centrally designed mould, the distinguishing mark of the new liberal eugenics is state neutrality" (ibid., p.

137). State neutrality is, of course, central to any liberal democratic system which aims to protect the principle of value pluralism. In its original formulation, the principle of value pluralism was primarily intended to safeguard freedom of religion and expression, but liberal eugenicists believe that it is broad enough to cover the freedom to use genetic technologies in the field of reproduction (Robertson, 1994). This means that governments must refrain from interfering not only with the more ordinary reproductive choices of its citizens but also with new reproductive choices made possible by genetic technologies.

Another important reason why liberal eugenicists are convinced that there is no need for moral panic in the face of a new eugenic era is that they think that there is no morally relevant difference between shaping humans by making modifications to their environment and shaping humans by making modifications to their genes (Agar, 1998, p. 137). They argue, for example, that if parents are allowed and even encouraged to increase their children's intelligence by providing them with the best possible education, then they should also be allowed to pursue the same goal through genetic technologies. There are two aspects to this claim. First, it allows liberal eugenicists to refute the common argument that genetic intervention is substantially more intrusive than any other influence we may have over the development of another human being. Second, if there is no substantial difference between genetic intervention and other influences that parents have over the development of their children, then there is also no need to develop new ethical guidelines and legal regulations for genetic technologies, because the freedom to use such technologies is already protected by the existing right to reproductive freedom (Harris, 2007, p. 75).

Yet, some critics have argued that there is indeed nothing morally suspect about human genetic modification as such but that one should nonetheless distinguish between genetic intervention for therapeutic purposes and intervention for enhancement (Walters & Palmer, 1997, p. xviii; Campbell et al, 1999, p. 76). There are two assumptions inherent to this argument. The first is that there is an objective difference between genetic interventions that aim at restoring the capacities of the body to their 'normal' state and interventions that aim at raising them above this state. The second assumption is that this distinction corresponds to the moral boundary between permissible and impermissible uses of genetic technologies. In other words, this argument holds that there is nothing morally wrong with using genetic technologies to

heal people, but that it is impermissible to use them to boost human capacities above what is normal, or for that matter, below what is normal (Scully, 2001).

While this argument appears to possess the merits of simplicity and fitness for practical application, both of its assumptions have met with severe criticism from liberal eugenicists. John Harris, for one, thinks that "enhancements are not plausibly defined relative to normalcy, to normal species functioning, nor to species-typical functioning" (Harris, 2007, p. 36). According to him, these notions "play no part in the definition of harm and therefore no part in the way the distinction between therapy and enhancement is drawn" (ibid., p. 46). He gives a striking example to illustrate this. Suppose it was possible to use genetic technologies to slow down or even halt the ageing process. If we would only allow genetic intervention to restore normal functioning, then we would have to forsake this clearly benevolent use of genetic technology because it would not simply restore our body to normal functioning but actually enhance it beyond its normal state. In other words, since it is perfectly normal for us to die of the diseases of old age, this intervention would go beyond the therapeutic use of genetic technologies and would therefore be morally unacceptable. As a libertarian consequentialist, Harris believes that the moral imperatives either to provide therapy or enhancement derive from the fact that we value minimizing harm and maximizing benefits. What counts in deciding if it would be permissible to use genetic technologies is not the fact of whether an individual's current state deviates from normal functioning, but the cost/benefit calculation regarding the body's "possible functioning" (ibid., p. 53). That is to say, the only pertinent questions are whether the harm the technologies aim to prevent is serious enough and whether the benefits they aim to produce are valuable enough to take the risks.

2. In Defense of Human Nature

Some critical liberal observers have argued, however, that the proposed marriage between eugenics and liberalism will not so much redeem the former of its authoritarian drift as corrupt the central principles of the latter to the point of its becoming something different altogether. This argument can take a variety of different forms, but the basic assumption is that modifications to the human genome threaten to disrupt something that is valuable in itself. Consequently, since our very understanding of human dignity and its legal reflection in human rights is founded upon the notion of human nature, then

genetic modification could ultimately signal the end of the central principles of liberal democracy. In The Future of Human Nature, Jürgen Habermas therefore argues that human nature should be legally protected against genetic enhancement. Yet he founds this claim on a very specific understanding of what it essentially means to be human. What he seeks to protect are not the 'species-typical' characteristics and behaviors of homo sapiens, but "the conditions under which the practical self-understanding of modernity may be preserved" (Habermas, 2003, p. 26). He argues that once we achieve a reflexive understanding of the necessary conditions for "our capacity to see ourselves as the authors of our own life histories" (ibid., p. 25), we will realize that a liberal eugenics contradicts these conditions and should therefore be rejected. Central to Habermas's argument is the notion that dignity is not a property one possesses simply by virtue of being human, but that it is the distinctive mode of being of a "communicatively structured form of life" (Ibid., p. 72). What he means by this is that we are only able to understand ourselves as free and autonomous agents worthy of respect in the context of a moral community that consists of equal members interacting with each other on the basis of norms and reasoning. Thus, when Habermas states that the danger of genetic technologies lies in their power to change human nature, he means that their free deployment threatens to undermine the very foundations of the moral community.

To demonstrate why this is so, Habermas invites us to consider that our lifeworld is still largely 'Aristotelian,' in the sense that we tend to make automatic distinctions between "what is manufactured and what has come to be by nature" (ibid., p. 46). This distinction is morally relevant insofar as it motivates us to adopt a particular mode of action when dealing with entities belonging to either one of these realms: while inert, inorganic entities are open to various forms of technical-instrumental intervention, self-regulated organic entities are not. According to Habermas, this is due to the fact that we spontaneously feel 'empathy' for organisms which seem to possess a certain amount of subjectivity, no matter how minimal. We remain committed to this logic in the case of genetic interventions carried out on embryos for therapeutic purposes, firstly because our actions in this case are still guided by the natural processes of growth inherent to this prenatal form of life, but also because we imagine how the future person might give consent for any intervention that could prevent or cure a debilitating condition. In the case of genetic enhancement, however, a very different scenario emerges. Here, prospective parents are not treating the embryo as another subject who will

come to be on an equal footing with them, but as an object they can simply dispose of if necessary. In other words, if the relationship between parent and child is reduced to that of producer and product, they will never be able to meet each other as equal members of the moral community.

According to Habermas, then, liberal eugenicists make the mistake of focusing solely on the freedom enacted in parental choice, while the proper question to ask is what consequences genetic intervention will have for the programmed person's "capacity of being oneself" (ibid., p. 57) on which one's ethical self-understanding as a free and autonomous member of a liberal egalitarian society depends. It is true that if "we experience our freedom with reference to something which, by its very nature, is not at our disposal" (ibid., p. 58). then the situation of the programmed person is not fundamentally different from that of an individual born "the natural way," for neither have had any say in the genetic traits and characteristics they are endowed with. The crucial question to ask, however, is if it makes any difference whether these traits are the result of natural chance or of the deliberate intervention of a third person. Liberal eugenicists tend to play down the impact of this intervention on the existential situation of an enhanced individual by suggesting that there is no substantial difference between improving a person by modifying her social environment and doing so by modifying her genes. In his view, however, while a genetically unenhanced person always retains the option of rejecting or reappraising her parents' attempts to shape her personality through socialization, the enhanced person "who is at odds with genetically fixed intentions is barred from developing (...) an attitude towards her talents (and handicaps) which implies a revised self-understanding and allows for a productive response to the initial situation" (ibid., p. 62). Moreover, a liberal eugenics would not only deprive the genetically enhanced person of the spontaneous self-perception of being the singular author of her own life, but also create the child's permanent and irreversible social dependence on the parent, which "is foreign to reciprocal and symmetrical relations of mutual recognition proper to a moral and legal community of free and equal persons" (ibid., p. 65).

3. The New Eugenics and the self-negation of Liberalism

It should be clear, then, that this debate mainly revolves around the question of whether the new eugenics concurs with or contradicts the central principles of

liberalism. Clearly, much depends on the actor that is given priority in this discussion: while liberal eugenicists tend to emphasize the parent, liberal critics believe that special consideration should be given to the prospective child. This difference of emphasis explains why the former group considers state neutrality in the domain of reproduction to be a sufficient guarantee of the liberality of the new eugenics. They argue that if the state remains neutral in this matter and does not intervene to enforce a particular conception of the good to be sought through genetic modification, then, by giving parents more control over which genetic traits their children will inherit, a liberal eugenics will actually strengthen the freedoms associated with reproduction. Critics, on the other hand, point out that the main threat to the central tenets liberalism no longer comes from potential state intervention but from parents themselves. According to this group, it is not the freedom of parents that is at issue but the freedom of the children born to them.

On the face of it, this way of framing the new eugenics debate may not be entirely satisfactory, for it gives the impression that liberal eugenicists believe there should be no moral or legal limitations whatsoever to the reproductive liberty of parents. This is obviously not the case. As with other individual liberties, reproductive choices tend to be judged for acceptability against John Stuart Mill's principle of harm. As is well known, this principle broadly states that one is free to act as one chooses, as long as one's actions do not cause harm to others. The problem in the specific case of genetic enhancement is, however, that the limit of individual freedom is not set by potential harm done to fellow citizens but to human beings who do not yet exist. One of the most influential approaches to this complicated issue was developed by Derek Parfit, and is known as the "nonidentity argument" (Parfit, 1984). The example Parfit gives is that of a 14-year-old girl who decides to have a child. Intuitively, we would be inclined to believe that she is likely to harm her child because, by dint of having such a young mother, the child is likely to receive "a bad start in life" (ibid., p. 358). Furthermore, we would probably also believe that it would have been better for her child if the mother had waited longer to conceive, for then her child would have had better chances in life. Parfit shows, however, that this is an inaccurate appraisal of the situation. If the girl had indeed waited longer to have a child, this child would have been the product of a different egg and a different sperm. It would, in other words, have been a different child. The further implication of this is that the child born to her at the age of 14 has not been harmed, since the condition of this particular child should be compared

not to that of the hypothetical child born a couple of years later but to the condition of not being born at all. In other words, being born to a 14-year-old mother is no worse for a child than being born to, for example, a 24-year-old mother, because the alternative is not being born at all. One of the conclusions that has been drawn from this argument is that in reproductive freedom, the threshold of harm should be set at the point where the child would have been better off not being born. The underlying rationale is that all forms of life which fall short of this threshold constitute a "life not worth living" or a "wrongful life" (Feinberg, 1986).

Obviously, the problem that some liberal eugenicists have with this argument is not that it would give prospective parents too little reproductive liberty, but that it would give them too much. Indeed, very few are willing to accept the ultimate conclusion to which this argument seems to lead, namely that parents' reproductive liberty should be so wide as to include even the freedom to endow their children with a physical or psychological disability. Yet, according to the nonidentity argument, a child would usually not be harmed by such an anomalous reproductive choice, for very few cases are likely to arise in which a child would find herself in such terrible conditions that it would have been better for her not to be born at all. It would, for example, be very difficult to maintain that being born deaf is worse than not being born at all. In order to escape this conclusion, liberal eugenicists usually fall back on what is called the principle of procreative beneficence, which, in one version, states that parents "should select the child, of the possible children they could have, who is expected to have the best life, or at least as good a life as the others, based on the relevant available information" (Savulescu, 2001, p. 413). Broadly, this principle entails that parents are morally required to give their children the best possible genetic endowment. It is clear, however, that this principle is still much too formal to prevent parents from endowing their children with a disability

How, then, do liberal eugenicists attempt to resolve this conflict between the principles of reproductive freedom and procreative beneficence in the case of selecting for disability? One solution could be, first, to define disability as a diseased state and subsequently argue that deliberately creating a disabled child constitutes a clear violation of medical deontology. This solution would not be wholly satisfying, though, because it would be necessary to reintroduce an objectivist notion of normality or normal functioning against which a given condition could be assessed. This is a solution that liberal eugenicists wish to

avoid at all costs. John Harris has therefore proposed defining disability as "a condition that someone has a strong rational preference not to be in and one that is moreover in some sense a harmed condition" (Harris, 2007, p. 91). To determine whether a given condition is a harmed one, he suggests using what he calls the "emergency room test:"

I have in mind the sort of condition for which if a patient presented with it unconscious in the emergency room of a hospital and the condition could be easily and immediately reversed, but not reversed unless the doctor acts without delay, a doctor would be negligent were she not to attempt reversal. (Ibid.)

According to Harris, the main advantage of this conception of a harmed condition is that it is not defined in relation to the state of nonexistence, or to normal functioning, but "relative to possible alternatives" (ibid., p. 92). Suppose, he explains, that someone was brought into the hospital with her little finger severed at the first joint and it could be sewn on again. Although it would obviously be absurd to maintain that the missing end joint of this person's little finger meant that her life would be not worth living, there are nonetheless good moral reasons to maintain that the hospital staff would harm the patient by failing to reattach the finger. According to Harris, the same holds true for all other injuries, diseases and disabilities.

Catherine Mills has fiercely criticized this definition of disability, firstly because it neglects the simple fact that "some disabilities are neither irreversible nor removable" (Mills, 2011, p. 22) and secondly because it uses the perspective of an "able-bodied person" (ibid.) as the standard against which to evaluate a given condition. Yet, though this criticism may certainly hold true in the present, Mills seems to ignore the fact that Harris develops this argument in relation to genetic modification technologies of the future. What he actually suggests is that when we have the choice to have a child either with or without a disability, we have good moral reasons to choose the second option. Another factor that critics have overlooked is that, as genetic science advances, it is likely not only to increase reproductive freedom and the responsibilities that come with it, but also to change the standards against which we seek to measure a harmed condition:

It is normal now, for example, to be protected against tetanus; the continued provision of such protection is not merely permissive. If the AIDS pandemic continues unabated and the only prospect, or the best prospect, for stemming its advance is the use of gene therapy to insert genes coding for antibodies to AIDS, I cannot think that it would be coherent to regard making available such

therapy as permissive rather than mandatory (Harris, 2007, p. 93; emphasis added).

We cannot think of a stronger argument against deliberately endowing one's children with a disability. What Harris is saying here is that if parents have the power to prevent their child being born with a disease or disability, they should have not merely the freedom to use this power, but "the obligation to pursue human enhancement" (ibid., p. 9) Harris's argument is not that the state should intervene to enforce this obligation—in his view, it is a moral obligation we have to our children—but we have no reason to assume that such demands will not be formulated as soon as these technologies become more widely available.

If, upon closer examination, the liberal eugenicists' argument for the freedom to intervene in the genetic make-up of future generations resembles an argument for the obligation to intervene, then it is seems that the critics are right to conclude that "liberal eugenics is a betrayal of liberal philosophy" (Fox, 2007, p. 24). Curiously, this is not how they themselves reach this conclusion. As we have showed, the danger that many see in a liberal eugenics is that it might change human nature. Habermas has developed what is probably the most sophisticated version of this approach. His main point of critique is that being endowed with specific genetic traits and characteristics will deprive the programmed person of "an unobstructed future of his own" (Habermas, 2003, p. 63). The idea is that a person who learns that some of her talents, skills and abilities were not given to her by "nature" but by means of the deliberate intervention of another person will find it impossible to understand herself as the singular author of her own life.

Interestingly, however, the underlying idea of this argument did not originate in the context of a discussion about the consequences of new genetic technologies. What actually prompted Joel Feinberg to write his seminal essay 'The Child's Right to an Open Future' (1980) was a series of lawsuits in which members of the Amish community challenged compulsory schooling laws in various states of the USA. As is widely known, the Amish live an extremely secluded life, far removed from the complexity of the modern industrialized world. In Wisconsin v. Yoder the United States Supreme Court ruled in favor of an appeal made by the Amish community, noting that by forcing Amish children to attend state schools the State of Wisconsin infringed on their constitutional religious rights. In his essay, Feinberg disagrees with this decision by arguing that the Amish way of life infringes on Amish children's

right to an open future by prematurely closing off many of the other ways of life available in a free liberal society.

It seems somewhat odd that Habermas refers directly to Feinberg's essay in his argument against enhancement technologies (Habermas, 2003, p. 124), for two reasons. Firstly, by likening attempts to shape children by altering their social environments to attempts to shape them by altering their genetic profiles, he actually seems to be pursuing one of the strategies that liberal eugenicists employ to argue the opposite of what Habermas himself intends. As explained above, if there are no substantial differences between genetic intervention and the other influences that parents have over the development of their children, then there is no reason to allow the latter while rejecting the former. Secondly, at the core of Habermas's argument lies the contention that while the effects of "a pathogenic socialization process" can always be "revised by critical reappraisal" (ibid., p. 62), this is impossible in cases of genetic intervention. If, as it appears, he actually disagrees with Feinberg's view on the intrusiveness of certain educational practices, why then does he claim to base his own argument on it?

The most plausible explanation for this confusion seems to be that Habermas wishes to retain the structure of Feinberg's reasoning but not its content. That is to say, he agrees with him insofar as we should be especially concerned about a child's right to an open future, but disagrees with him insofar as he rejects the notion that the greatest threat to this right comes from a "pathological" socialization process. Liberal eugenicists often liken the effects of socialization to those of genetic intervention in order to argue that the latter is no more problematic than practices that are now routinely accepted as part of normal parenting. Habermas would be unlikely to disagree with the argument that parents' reproductive freedom should also encompass genetic interventions. As soon as priority is given to the perspective of the 'passive receiver,' however, then a very different picture emerges. After all, whereas socialization occurs at a moment when a child is already able to respond to the actions of her educators, genetic intervention occurs before the child has even entered into existence and the resulting individual will therefore be unable to respond effectively to his or her producer's intentions:

> (...) such an imposition from within the community, even if it is excluded from the relationships obtaining between morally acting persons, must nevertheless not be confused with an external or alien determination of the natural and mental constitution of a future person, prior to an entry into the moral community (ibid., p. 79).

Most critics revert to the notion of human dignity in order to oppose these kinds of interventions. However, this road is not open to Habermas because it entails giving full rights to unborn life and in his view, it is clear that the question of whether "the in vitro embryo were already 'another,' who possessed completely valid basic rights (...) can hardly be answered in the positive given the premises of an ethically neutral constitutional order" (ibid., p. 77). If it is already extremely difficult—if not impossible—to reach consensus on the question of when life begins, then these problems are only likely to increase in the case of genetic intervention, for gene modification can be performed not only at the zygote and embryo stages, but also in sperm and egg cells. It seem quite reasonable to assume that few would be willing to accept the absurd consequences that would follow from giving sperm and egg cells full human rights. While the proposed dilemma is quite clear, however, the same cannot be said about Habermas's solution to it. When he contends that "legal protection might come to be expressed in a right to a genetic inheritance immune from artificial intervention" (ibid., p. 27), then it remains far from evident who might be the beneficiary of this right. Since he rejects the idea of giving such a right to prenatal forms of life, he seems to mean that it would be bestowed upon the adult enhanced person. But how could such a person ever exercise her right to a genetic endowment free from artificial intervention, given that this irreversible act would have taken place well before she was a position to do so?

There is more to be said here, though. For what the debate between the advocates and opponents of a liberal eugenics makes evident is that the emergence of enhancement technologies is likely to be accompanied by a growing tendency to impose severe normative constraints on certain potential forms of life. This is clear enough in the argument of someone like Habermas, who draws on a normative conception of human nature to argue against genetic enhancement. What has not been sufficiently emphasized thus far, however, is the fact that any attempt to give normative content to human nature may be mobilized politically to exclude those who deviate from this norm (Mendieta, 2003). That is not to say that these authors' conceptions of human nature could serve as grounds for excluding certain vulnerable groups, such as the disabled or the mentally ill, from the moral community. Instead, it could be said that these definitions preemptively deny any genetically enhanced being that may be brought into existence in the future the status of human being. What else could Habermas mean when he writes that "[t]his new type of relationship

[between programming parent and programmed child, author's remarks] offends our moral sensibility because it constitutes a foreign body in the legally institutionalized relations of recognition in modern societies" (Habermas, 2003, p. 14; emphasis added)? This statement seems, moreover, to cast further doubt on the effectiveness of Habermas's call for a right to a genetic constitution free from genetic intervention. If an enhanced person is barred from establishing reciprocal relationships with 'normal' human beings, and thus from entering the moral community of equal citizens, on what grounds, then, may such a person appeal to this right in the first place? Again we must ask who the bearers of this right would be if the only individuals to have an interest in it were denied legal subjectivity?

This tendency is not absent from the discourse of liberal eugenics, however. Quite the contrary, in fact. As explained above, many liberal eugenicists seek to avoid some of the more distressing consequences of the nonidentity argument by tempering the right to reproductive liberty with the principle of procreative beneficence. The ultimate result of this argument is that parents would have the obligation both to prevent their children from being born with a disability or with a disease and to boost their capacities to a maximum. We should not lose sight of the justification behind this line of reasoning, however. What liberal eugenicists reject is not the notion of 'wrongful life' as such, but only the criteria which are to be used to determine what forms of life are included in this category once genetic technologies become available. What they are actually arguing, therefore, is that while it may be true that it is currently better, for example, to be born deaf than not born at all, this may change once we have the power to choose between a deaf child and a hearing child. If it is true, on the other hand, that the emergence of genetic technologies will progressively raise the threshold of harm, then we are also about to witness a steady increase in the number of forms of human existence that will have to be categorized as wrongful life. It remains to be seen how far this category can be stretched but perhaps, in the not too distant future, human beings as we currently know them will all be judged as having a 'life not worth living.'

Conclusion

Human genetic modification is still in its infancy, but the issues discussed above suggest that liberal political and moral philosophy remains rather ill equipped to address this controversial field, in the sense that the two positions

appear to be conceptually inconsistent: ultimately, both lead to conspicuously illiberal conclusions. After all, as we have argued above, while the argument against a new eugenics necessarily entails a preemptive dehumanization of any enhanced form of life, the argument for it threatens to reduce any non-enhanced form of life to the status of wrongful life. The final analysis might conclude, then, that any kind of liberal response to the challenges of the new eugenics unwittingly produces a form of life devoid of any intrinsic value. This is not to say that this outcome is inevitable, but clearly we will need to rely on an alternate interpretative framework if we wish to gain a more precise understanding of this contentious issue.

REFERENCES

Agar, N. (1998). Liberal eugenics. *Public Affairs Quarterly, 12*, 137–155.

Agar, N. (2004) *Liberal Eugenics: In Defence of Human Enhancement.* Oxford: Blackwell.

Campbell, A., Gillet, G., & Jones, D.G. (1999). *Medical Ethics*, 2nd edition. Oxford: Oxford University Press.

Dworkin, R. (2000). *Sovereign Virtue: The Theory and Practice of Equality.* Cambridge, MA: Harvard University Press.

Feinberg, J. (1980). The Child's Right to an Open Future. In Aiken, W., & LaFollette, H. (eds.), *Whose Child? Children's Rights, Parental Authority, and State Power.* Totowa: Rowan & Littlefield.

Feinberg, J. (1986). Wrongful life and the counterfactual element in harming. *Social Philosophy and Policy, 4*, 145–178.

Fox, D. (2007). The Illiberality of "liberal eugenics". *Ratio, 10*, 1–25.

Fukuyama, F. (2002). *Our Posthuman Future. Consequences of the Biotechnological Revolution.* London: Profile Books.

Galton, F. (1883). *Inquiries Into Human Faculty and its Development.* London: J. M. Dent & Company.

Habermas, J. (2003). *The Future of Human Nature.* Cambridge: Polity Press.

Harris, J. (2007). *Enhancing Evolution: The Ethical Case for Making Better People*. Princeton: Princeton University Press.

Mendieta, E. (2003). Communicative freedom and genetic engineering. *Logos, 2*, 124–139.

Mills, C. (2011). *Futures of Reproduction. Bioethics and Biopolitics*. Dordrecht: Springer.

Parfit, D. (1984). *Reasons and Persons*. Oxford: Clarendon Press.

Paul, D.P. (1992). Eugenic Anxieties, Social Realities, and Political Choices. *Social Research, 59*, 663–683.

Robertson, J.A. (1994). *Children of Choice: Freedom and the New Reproductive Technologies*. Princeton: Princeton University Press.

Sandel, M.J. (2008). *The Case Against Perfection. Ethics in the Age of Genetic Engineering*. Cambridge, MA: Harvard University Press.

Savulescu, J. (2001). Procreative beneficence: why we should select the best children. *Bioethics, 15*, 413–426.

Scully, J.L. (2001). Drawing a line: Situating moral boundaries in genetic medicine. *Bioethics, 15*, 189–204.

Stock, G. (2002). *Redesigning Humans: Our Inevitable Genetic Future*. Boston/New York: Mariner Books.

Walters, L., & Palmer, J.G. (1997). *The Ethics of Human Gene Therapy*. New York: Oxford University Press.

Wilkinson, S. (2008). "Eugenics talk" and the language of bioethics. *Journal of Medical Ethics, 34*, 467–471.

Towards a "Human Enhancement Society"?
Opportunities for an Aristotelian Approach
to Frame the Question

Alberto Pirni [†]
a.pirni@sssup.it

ABSTRACT

The essay is subdivided into three parts. In the first and introductory one the current debate on human enhancement is presented, with specific reference to its interdisciplinary characteristics and to the aspects which explicitly challenge "the human condition" as a whole. The second and third parts attempt to frame the comprehensive area of questioning opened by such a perspective, which is grounded in the practical philosophy of Aristotle – a model that seems particularly neglected within the human enhancement debate. Specifically, part two (§ 1) is devoted to a "rehabilitation" of the theory of justice and fairness developed in the *Nicomachean Ethics*. In turn, part three (§ 2) goes into detail with reference to the taxonomy used, and tries to sketch out a possible area of theoretical application regarding both the rights of restoring and possible criteria of legitimate advantage. The proposal outlined is also integrated by a synthetic list of possible points of criticism which might be taken seriously into account in a wider and deeper exploration of this approach to the topic.

Introduction

The debate on human nature – that is *why* man exists as such, on *how* he has been able to evolve, and on *who* he can become in the near or more distant future – constitutes a combination of questions that are constantly and inexhaustibly proposed.[1] This is perhaps the outcome of an intrinsic specificity

---

[†]Scuola Superiore Sant'Anna – Pisa, Italy.
[1]For a recent book on the topic, see Downes & Machery (2013).

of human beings. As has been acutely observed, in no other living species does technology (or more precisely the need for technological development), play such an important role as it does in the human species (Gerhardt, 2008, pp. 98 ff.). Almost every animal species is able to implement "techniques" in order to improve its living conditions, that is to produce tools or structures capable of enhancing its ability to procure food and defend itself. No animal, however, to the same extent as the human being, seems unable to do without the constant innovation of its own capacities and their outcomes. This is a constant factor which unequivocally connotes our being human. It is a constant that we can say is expression of an essentially dual nature: fear of inadequacy, on the one hand, anxiety for perfection, on the other.

It is perhaps this very same constant which is currently active in one of the most virtuously interdisciplinary debates that is taking place on the contemporary scene and which is frequently labelled as human enhancement. The expression "human enhancement" at the beginning of this debate substantially had the meaning of an intervention cogitated in order "to develop the appearance or functioning of the human being beyond what was necessary to sustain and re-establish good health" (Juengst, 1998, p. 29). In definitions like this, of a bioethical form and to this day particularly influential, at the centre of attention there is the binomial of illness/health. But the debate has developed rapidly, introducing different definitions of the very concept of human enhancement which, together with that of "health", imply the concept of "normal functioning" but also those of "therapy", "well-being", "dignity" and so on. A particular characterization of this debate developed throughout the first decade of the 21st century, polarizing along two aspects of the argument, often presented as ideologically opposed, the *transhumanist* viewpoint and the *bioconservative* viewpoint.[2] Very briefly, those who follow the first line of argument, sees the current form of the human species, on both a somatic and cognitive level, as constituting only one of the stages of human development, whilst we have only just begun to grasp the universe of possible integrations between natural and artificial that the succeeding phase of this development will involve. Conversely, the other line of argument stresses the need to investigate the significance and implications of the transformations concealed behind the apparently neutral technological development involving

---

[2] There is an attentive and equidistant comparison in Parens (1998). For a more updated framework, see Carnevale & Battaglia (2013).

the human subject, framing the concepts of nature and human dignity as insurmountable limits. Whilst the exponents of the first line of reasoning are criticized for being entrenched in an implicit (and naive) determinism regarding the progressive unproblematic development of the human species, those of the second line of argument are chastised for their excessive "metaphysical" vagueness regarding the basic concepts introduced in defence of more cautious positions.

The main merit of this development of the debate can be found in a more widespread knowledge of the underlying implications regarding his issue and a series of interventions on the human body that make invasiveness, non-reversibility and growing technological integration its distinctive characteristics. An indication of this awareness has been and still is the increasing entrustment, on the part of public bodies and institutions, of research aimed at better mapping a phenomenon that is difficult to contain and which has potentially unpredictable outcomes. The clearest result, following an important series of documents regarding this research, are definitions that are decidedly "wider", or better "more comprehensive" of the various possible aspects of the phenomenon. A significant example is that of the work commissioned by the European Parliament and involving the research group coordinated by C. Coenen. Here, human enhancement is seen as "a modification aimed at improving individual human performances and determined by interventions carried out on a scientific or technological basis on the human body" (Coenen, 2009, p. 17).[3]

It is clear, from definitions like this, that there is awareness of the fact that the issue of human enhancement goes very much beyond the borders of the relationship illness/therapy. At the point where this label is used for high value technological interventions ranging from aesthetic surgery to pre-implant genetic diagnosis, from empowering chemical compounds that determine enhanced performances to bionic prostheses or wearable exoskeletons, there appears to be much more than the medical sphere involved. There is the perception that the entire human condition comes into play, to recall the famous expression of Arendt in a sense perhaps not yet unveiled to the German philosopher.

---

[3] The document, elaborated by the research unit *Science and Technology Options Assessment* (STOA), a part of the European Parliament, can be found at: http://www.europarl.europa.eu/RegData/-etudes/etudes/join/2009/417483/IPOL-JOIN_ET(2009)417483_EN.pdf.

The profound refection that such a radical upheaval will cause in the near future must depart from some basic assumptions that are already transversally well-known. Above all, it appears that we can reject monistic viewpoints, those in absolute agreement or absolute disagreement as regards this set of modifications, of unprecedented complexity and extremely rapid quali-quantitative growth. Secondly, it does not seem that the understanding of such an amalgam can be entrusted to a single analytical perspective, but certainly requires a decidedly interdisciplinary approach (Straub, 2012; Grion, 2012; Cerqui, 2002).[4] Thirdly, it appears to be a particularly demanding task to identify precisely the very object of this approach, that is distinguish a clear "set" of technologies explicitly oriented towards human enhancement, given the highly diverging and incomparable characteristics they represent, from intervention on single cell groups to structures or bio-robotic prostheses, from neurological interactions with external computers on the single subject to tools involving the sensorial enhancement of touch or sight, to cite but a few examples.

With respect to all this, together with a pragmatic attitude, contextualized and as free as possible from ideological implications, such as that recently inaugurated by the U.S. National Science Foundation (Allhoff, 2009), it is the intent in that which follows, inevitably in a preliminary form, to put a different theoretical-critical modality to the test. The need for an anti-monistic and anti-reductionist approach, together with the need to appreciate the contextual conditions in which this enhancement should be placed and with the need to examine case by case with its risks and opportunities, in fact make the attitude appear to be fronetic and the entire perspective of Aristotelian practical philosophy as a potentially useful orientative tool and currently ignored by the present debate on this issue. Furthermore, this perspective seems to connect with one of the developments of a pragmatic nature regarding this theme which appear today to be most fruitful but, equally, are not totally aware of the preciousness of the Greek antecedent which can equip a toolkit suited to contemporary complexity.[5]

Therefore, in what follows there is the intent to launch an exploration of a scheme of Aristotelian matrix for different possible applications to ethical-political questions relative to what we could term the "human enhancement

---

[4] This consideration has also been made by some that the extreme interdisciplinary nature of an approach suitable for human enhancement urges the need for a totally new discipline (Savulescu and Bostrom, 2009).

[5] I refer above all to the commendable volume of Keulartz et. al. (2002).

society". We are certainly speaking of a futuristic perspective, but which, in many respects, already appear real and looming over the present.

1. Towards An Aristotelian Approach I: Human Enhancement through the Lens of the Theory of Justice

In order to try and elaborate a scheme of this type, it seems opportune to return to the heart of Aristotelian practical philosophy and, in particular, to the virtue to which the Greek philosopher dedicates a wholly exclusive space within the context of the *Nichomachean Ethics*: *justice*.[6]

1.1. What is justice? The first Aristotelian answer

Here, as we know, Aristotle starts by considering justice as virtue ethics, on the one hand. On the other hand, justice is immediately treated as a very special virtue: justice is considered as the "complete virtue in its fullest sense". Following Aristotle's words:

> This form of justice, then, is complete virtue, but not absolutely, but in relation to our neighbour. [...] It is complete because he who possesses it can exercise his virtue not only in himself but towards his neighbour also; for many men can exercise virtue in their own affairs, but not in their relations to their neighbour. [...] Now the worst man is he who exercises his wickedness both towards himself and towards his friends, and the best man is not he who exercises his virtue towards himself but he who exercises it towards another; for this is a difficult task. (Eth. Nic., 1129a 25 – 1130a 8)

A crucial point stressed here by the author is that the perfection of justice does not come so much from its being a virtue (that is "a moving towards the good") of the person who exercises it, but rather from his being engaged in seeking the good of others. Thus, according to a first approximation, what does justice mean? Aristotle proposes here two meanings, one wider: "The just, then, is the lawful", and a second and narrower one (as we will discover later on): "The just is the fair" (Eth. Nic., 1128b 34). By expressing the latter meaning in other terms we could rephrase it as: "just is what respects equality". We must remember that Greek does not have a vocabulary of fairness distinct from that of equality. And this becomes evident by looking at the conduct of an unjust

---

[6] The English Translation by D. Ross of the *NichomacheanEthics* will be used in what follows, as well as the reference to the traditional pagination.

man who – says Aristotle – is a man "grasping" (Nic. Eth., 1129b 5), d.i. a man who seeks to obtain more (in terms of goods) than others. But this is not sufficient (even if Aristotle is not so explicit here about this): the unjust seeks to obtain more than others "by grasping", namely: through unfair means, or by using unfair or undeclared forms of advantage.

1.2. Which Justice? A Second Definatory Frame

Following the argumentative path of the fifth Book, Aristotle appears as mainly (even if not exclusively) committed to the definition of this second and narrower sense of justice: justice as respect of equality. And this is the core assumption of the argumentation presented in paragraphs 2-4 of this Book. In this context the Greek philosopher puts forward the paradigmatic distinction between *distributive* and *rectificatory justice* (Nic. Eth., 1130b 30 – 1131a 1). Thus, *distributive justice* concerns only public goods, namely public offices, which – this is the line which will be developed in the following paragraphs – should be distributed only by respecting a ranking of merits and of relationships of proportionality among the community's members (to give to each one what he/she deserves proportionally to their role within the polis). This is the particular way of interpreting the claim of equality in the public domain.

Conversely, there also exists a second form of justice: *rectificatory justice*; this is always related to the restoring of equality, but in this case concerning private relationships, namely, relations among "privates". This is a quite generic category put forward for identifying single citizens first, but also, more generally, human beings (including in this way also women), or beings that cannot be completely considered as human (slaves) or that are not yet fully human (children). Therefore, rectificatory justice appears crucially engaged in the interpretation of that relationship between individual and otherness, which justice as a virtue represents in an exclusive manner. In paragraph 4 the author defines more sharply the concept of rectificatory justice:

> justice in transactions between man and man is a sort of equality indeed [...]. For it makes no difference whether a good man has defrauded a bad man or a bad man a good one, nor whether it is a good or a bad man that has committed adultery; the law looks only to the distinctive character of the injury, and treats the parties as equal. (Nic. Eth., 1131b 32 – 1133a 5)

Such justice is a sort of equality, but the author does not forget that justice is also an ethical virtue. As virtue, justice constitutes firstly as a *metriotes*, a sort

of mediation between two extremes, namely between a *maximum* and a *minimum* (Nic. Eth., 1131b 32 – 1133a 5.).Thus, justice is a medium point between a sort of gain and a sort of loss; "it consists in having an equal amount before and after the transaction" (Nic. Eth., 1132b 19-20). The very distinctive point is its orientation to the restoring recovery from an existing inequality, or a disparity, a dis-equilibrium – innate or which is formed at a certain point – between a gain and a loss.

1.3. Beyond Justice: The Role of Fairness

Chapter 5 of the *Nicomachean Ethics* offers an area of questioning which is certainly richer than can be taken into account in the present schematic context. However, simply to give an example contemplated within Aristotelian theory, a "complete" theory of justice, which is aimed at grasping the problematic whole presented by the technological age, will surely be faced with the problem of "justice and liability" or with the question as to whether "it is possible to suffer injustice voluntarily".

Among other concepts and arguments, almost at the end of the Book, Aristotle introduces the concept of *fairness*, which we have to consider analytically. The author devotes his preliminary attention to a fundamental statement clearly aimed at inscribing fairness within the framework of justice.

For on examination they appear to be neither absolutely the same nor generically different; and while we sometime praise what is fairness and the fair man [...] at other times, when we reason it out, it seems strange if the fair being something different from the just, is yet praiseworthy; [...] they are all in a sense correct and not opposed to one another; for the fair, though it is better than one kind of justice, yet is just, and it is not as being a different class of thing that it is better than the just (Nic. Eth., 1136b 31 – 1137a 5).

Fair is part of the same class of things and of being of the just. Nonetheless, it is in a sense superior: "it is better than the just". Aristotle stresses this point immediately after, integrating it with additional defining elements.

The same thing, then, is just and fair, and while both are good the fair is superior. What creates the problem is that the fair is just, but not the legally just but a correction of legal justice. (Nic. Eth., 1137b 10-13). But, one could ask, why is the fair a correction and why and in which sense does the just need a correction? The reason for this is intrinsic to the same nature of the law:

all law is universal but about some things it is not possible to make a universal

statement which shall be correct. [...] When the law speaks universally, then, and a case arises on it which is not covered by the universal statement, then it is right, where the legislator fails us and has erred by oversimplicity, to correct the omission – to say what the legislator himself would have said had he been present, and would have put into his law if he had known. Hence the fair is just, and better than one kind of justice – not better than absolute justice but better than the error that arises from the absoluteness of the statement. And this is the nature of the fair, a correction of law where it is defective owing to its universality. (Nic. Eth., 1137b 14-27).

It can occur that the universality of the law, due to a particular case (a new one which is not yet contemplated in that law), must be corrected by the intervention of a sort of "second legislator", the fair man, who seeks to adapt the universality of the statements of a law to the cases that may not fit it, or that could be also deeply misunderstood, if they were not submitted to such an intervention of fairness. Along this path, we should note the terms in which Aristotle finally depicts the fair man: he is "the man who chooses and does such acts, and is no stickler for his rights in a bad sense but tends to take less than his share though he has the law oft his side" (Nic. Eth., 1137b 14-27).

The fair man is not the man who abuses from his position of (indirect) prominence in order to gain advantages for himself or for the members of his restricted community. On the contrary, the fair man is he who takes less than what the law would have allowed and assigned to him.

Restoring the previous schema, which we saw in relation to rectificatory justice, we could claim that the fair man, instead of putting himself on the side of advantage, prefers to occupy the side of loss, in order to give more place and possibilities to a wider community of possible others, in the present and for the future. He takes less for himself, in order to leave more for others: this is perhaps the best and deepest spirit of justice in an embodied form that the fair man represents.

2. Towards an Aristotelian Approach II: Opportunities, Limits, Open Questions

It is unfortunately not possible within this context to fully elaborate all the points that in the fifth Book of the *Nichomachean Ethics* returns, perhaps a little unexpectedly, to the reader interested in its possible application to today's technological society. In what follows there is the intent merely to trace the profile of a possible taxonomy of justice and fairness of an Aristotelian matrix to put in relation to the forms of human enhancement, limiting oneself

to evoking some problematic aspects and theoretical nuclei that certainly deserve a succeeding and more analytical close examination.

2.1. Going Back to Aristotle – I

Firstly, as can be recalled, justice consists in the search for good not only from the viewpoint of the first person, but also, and above all, "for others". Furthermore, since the unjust man is first of all he who attempts to obtain more of others "by grasping", that is to say: using disloyal or undeclared forms of advantage, it will be corrective justice that will identify and re-establish an intermediate way between a sort of profit and loss. Its specific difference is in fact its orientations towards recovery, the restoration of an existing inequality, or from a disparity, a dis-equilibrium– we could say: innate or which has been generated at a certain point – between a profit and a loss, in all the terms in which both can be possibly imagined.

From this point of view, one could put forward here a first proposal of development of the model of corrective justice, directing it explicitly towards all the needs of recovery from situations of damage or disability, congenital or acquired following traumatic events, illnesses or aging. In these cases human enhancement can be treated as a form of profit relative to the recovery from a form of loss. Extensively, we could call the questions that are a part of this context "questions of justice" for a human enhancement society.

These questions should be aimed – as far as is possible today, thanks to the multi-formed technological contribution – at restoring capacity and functioning typical of a "healthy" adult person to that person who has lost these capacities and functioning or has never possessed them.[7]

Starting from an analytical elaboration of all the public questions that could be part of this perspective might take the first steps towards a new system of public policies, committed to considering the set of technological innovations directed towards human enhancement as a possibility of rehabilitation of high standards of quality of life for each and every citizen of a community. We should certainly analytically consider each context in which each possible technological innovation could evoke claims of corrective justice in the sense clarified above. The field of the new generation of biomedical applications – such as bionic hybrid systems, bio-mechatronic prostheses and components for

---

[7] Here the reference is to the terms *capacity* and functioning in the meaning elaborated in Nussbaum and Sen (1993).

sensorial and motor augmentation – but also the cutting edge context of biomedical research – consider the universe of nanotechnologies or neural interfaces – certainly presents problematic areas that require new and specific languages of justice. If one retains that there are margins of fruitfulness, this is certainly a pathway to proceed with entirely *ex novo*.

2.2. Going back to Aristotle – II

A suitable integration of this provisional and partial scheme is without doubt offered by the reference to fairness. As we can recall, the fair is something superior to the legal just and constitutes a sort of corrective to the latter in relation to single cases. Specifically, the just man aims to obtain fewer advantages for himself than that which the law would have allowed. In this way, let us say, the just man operates a correction, individual and voluntary, to the advantage of possible others.

Bringing this problematic combination back to the context of human enhancement, in a first approximation one could establish the criterion on the basis of which the enhancement of the same sphere of human capacities and functionings beyond the line of what is "normally" attributed to a "healthy" adult person is considered as a form of advantage that must be compensated with a corresponding forma of loss or of "restoration" of the positive consequences of these advantages to his community of reference.[8]

This criterion cannot avoid some clear boundaries: first of all, we should preserve the maximum space for the self-determination of the individual, where the enhancement required does not infringe existing laws.[9] However, if this determination has directly or indirectly an impact on others, producing situations of undesired and suffered disadvantage, we must have the political-juridical possibility of arresting this self-determination and assessing its possible effects under the lens of fairness. Corresponding to this view, there could depart from here a hypothetical list of "questions of fairness" for the human enhancement society.

The basic objective of these questions should be the elaboration of criterion of legitimate advantage: not all that which can be done must also be realized

---

[8] On this theme, for a first framing of a very wide question, see again Lucivero & Vedder (2013).

[9] The theme of regulating spaces of legitimate spaces of freedom with respect to the challenges posed by new technologies is certainly one of the main debating points within this context. For a preliminary framing see the volume Palmerini & Stradella (2013) and, in it, the essay Pirni & Carnevale (2013a).

(see for example the voluntary and informed use of doping by an athlete, but also a possible neural implant which emphasizes unexpectedly my cognitive or memory capacities). Pandering acritically to the possibilities of implementation of our body and brain put at the disposal of the advancement of technology can open the door to also relevant distortions of social cohesion and free competition between individuals within a community. The risk of damage that is difficult to evaluate diachronically seems clear: an individual advantage could reveal itself to be a social loss and an undeserved human improvement could trigger chains of iniquity difficult to compensate.

2.3. Mapping a Territory of Open Questions

Wishing to attempt a summary of the profiles of this taxonomy, one could affirm that, whereas a theory of justice (in terms of corrective justice) suited to the challenges of human enhancement should be aimed at the legitimate protection of the recovery of loss or injury, a parallel theory of fairness should have the objective of safeguarding the legitimate improvement.

Obviously this articulation seems still very much preliminary and needful of integration. One could add as support that a first and perhaps most urgent task for an ethical and political theory which is directed towards this context of reflection is that of offering a mapping of the territory of open questions which is as analytical and detailed as possible – often in a radical manner – from the combination of possibilities and risks offered by the various dimensions of human enhancement. However this cannot, already at this level, avoid the emersion of a preliminary series of objections.

With respect to the aspect of corrective justice, whilst it appears quite clear that an intermediate point can be identified for the metriotes between "normality" and "disability" at a physical level (for example in the case of a mutilation or evident limitation of a limb), this could be much more complex to do in the case of cognitive or psychic damage in terms of the status of what is "healthy".

Conversely, with respect to the aspect of fairness, it would seem clear to have the right to interrupt "egoistic" forms of distortion and improvement. But are the questions brought up above in this regard "juridifiable", that is able to launch a legitimate legislative production within a democratic juridical framework?

Again, under a more general ethical-juridical profile, one could wonder whether and to what extent the current legislation regarding privacy, responsibility and informed consent is adequate, with respect to the long-term effects of devices that are often available for use even if they are beyond the ordinary standards of reliability. One should also discuss the issue of the risk of non-reversibility, together with that of the definition of the duration and legitimate aims of these "enhancements". Last but not least, one should understand the degree of exclusivity or inclusivity of these procedures, with the aim of avoiding or at least contemplating ex ante new and more or less explicit forms of "divides", subtler and more specific than the well-known "digital divide", that is the insurmountable differences between who can afford and not afford technological enhancements (Pirni & Lucivero, 2013).

Conclusion

As regards what has been presented so far, whilst on the one hand it has the intent of suggesting a line of framing of Aristotelian matrix of questions of justice and fairness for the technological society, on the other hand it prompts detailed investigation of the foundations of problematic contexts, like those now presented only summarily which, in reflecting the objective difficulty of the questions regarding human enhancement, make an answer to these questions still more urgent.

REFERENCES

Allhoff, F. et al. (2009). *Ethics of Human Enhancement: 25 Questions and Answers.* Washington, DC: US NationalScience Foundation Report.

Aristotle (2009).*The Nicomachean Ethics.*Transl. by D. Ross (or. ed. 1954); revis., introd. and notes by L. Brown. Oxford: Oxford University Press.

Carnevale, A., & Battaglia, F. (2013). A 'Reflexive' Approach to HumanEnhancement: Some Philosophical Considerations, in Lucivero &Vedder (2013), 95-116.

Cerqui, D. (2002). The Future of Humankind in the Era of Human and Computer Hybridization: An Anthropological Analysis. *Ethics and Information Technology*, *2*(4), 101-108.

Coenen, C. et al. (2009). Human Enhancement, Brussels: European Parliament 2009 (available at: http://www.europarl.europa.eu/-RegData/etudes/etudes/join/2009/417483/IPOLJOIN_ET(2009)417483_EN.pdf.)

Downes, S.M., & Machery E. (eds.) (2013). *Arguing About Human Nature. Contemporary Debates.* New York: Routledge.

Gerhardt, V. (2008). Homo publicus. In Ganten, D. et al. (ed.), *Was ist der Mensch?* Berlin/New York: de Gruyter, 97-102.

Grion, L. (ed.) (2012). *La sfida postumanista. Dialoghi sul significato della tecnica.* Bologna: il Mulino.

Juengst, E.T. (1998). What does enhancement means?, In Parens (1998), 29-47.

Keulartz, J., et al. (2002). *Pragmatist Ethics for a Technological Culture.* Dordrecht/Boston: Kluwer Academic Publishers.

Lucivero, F., Vedder, A. (eds.) (2013). *Beyond Therapy v. Enhancement? Multidisciplinary Analyses of a Heated Debate.* Pisa: Pisa University Press.

Nussbaum, M.C., Sen, A. (eds.) (1993). *The Quality of Life*, Oxford, Clarendon Press.

Palmerini E., Stradella E. (eds.) (2013). *Law and Technology. The Challenge of Regulating Technological Development.* Pisa: Pisa University Press.

Parens, E. (ed.) (1998). *Enhancing Human Traits. Ethical and Social Implications.* Washington, DC: Georgetown University Press.

Pirni, A., & Carnevale, A. (2013a). The Challenge of Regulating Emerging Technologies. A Philosophical Framework. In Palmerini & Stradella (2013), 59-75.

Pirni, A., & Carnevale, A.(eds.) (2013b). Robotics and Public Issues: a Philosophical Agenda in Progress. *Cosmopolis*, *2*(8).

Pirni A., & Lucivero, F. (2013). The 'Robotic Divide' and the Framework of Recognition: Re-articulating the Question of Fair Access to Robotic Technologies. *Law, Innovation and Technology*, *2*(5), 147-171.

Savulescu, J., & Bostrom, N. (eds.) (2009). *Human Enhancement.* Oxford: Oxford University Press.

Straub, J. (ed.) (2008). *Der sich selbst verwirklichende Mensch.* Bielefeld: transcript.

Human Presence and Robotic Mediations: An Alternative Framework for Explicating Human Enhancement[*]

Pericle Salvini [†]
p.salvini@sssup.it

ABSTRACT

In this essay I propose an alternative theoretical framework for explicating human enhancement. The framework is based on the concepts of reciprocity, which I consider a fundamental aspect of human presence, and of mediation, which I consider a fundamental aspect of the relation between human beings and science and technology. I argue that enhancement is given by the way in which technological and scientific mediation alters the structure of the network of reciprocity characterising human presence. As a matter of fact, technological mediation may turn the reciprocity of presence into a unilateral relation, which prevents any form of response. The lack of responsibility, here understood as the possibility to respond back, is determinant for the generation of a situation of power and for eliciting moral disengagement. The framework will be applied and discussed with reference to robotics technologies.

> In my opinion, there cannot be progress in the field of technology unless by means of criticism. We cannot be interested in a technological product unless we are interested in its negativity
>
> (Virilio, 1995).

1. Introduction

Is technological and scientific evolution an unavoidable aspect of human nature? With respect to all animal species, human beings are paradoxically the less fitted for surviving in the natural environment, since they lack a specific instinctual baggage. Nevertheless, the lack of instinctual capabilities is compensated by technological and scientific actions, which are a distinctive feature of the human species, turning human beings into the most adaptable and powerful of all living animals (Gehlen, 1988 [1940]).

From stones to computers and now robotics, the primary and original function of technological and scientific development has been to provide human beings with the tools and knowledge necessary to tame and exploit the natural environment as well as to protect and improve the frail nature of the physical body. As a matter of fact, while science is usually described as concerned with understanding the natural world, technology is usually described as 'the innovation, change or modification of the natural environment, in order to satisfy perceived human wants and needs' (ITEA, 2000/2002).[1] However, even if the contrary were true, it would have been impossible to prove it today.

Whatever it is the relationship between human beings and technological and scientific development, with respect to the time in which the wheel was invented and the fire discovered, it is unquestionable that the natural environment has now become increasingly artificial, that is, pervaded by cultural artefacts. Moreover, technological and scientific advancements have started to produce negative consequences on human beings and on the natural environment (e.g. pollution). Some of these negative effects are even independent on how well or bad we use technology (e.g. internet addiction). To say it with John Culkin 'we shape our tools and thereafter they shape us' (Culkin, 1968). In other words, it seems that there has been a reversal between goals and means: science and technology are no more needed only for mastering the natural environment and protecting (and extending the limits of) our human body, but they have become indispensable to master the complexity of scientific and technological development itself and repair to the problems it causes (Galimberti, 1999).

---

[1] Taking into account current scientific research trends, in particular in the field of biotechnologies, I would update this definition by including "the human body" among the targets of technological innovation, change and modification.

Due to the above and several other reasons, such as the almost unlimited possibilities offered by science and technology and the relevant role played by economic interests in steering scientific and technological objectives,[2] the need to find a limit and a meaning to technological and scientific evolution is becoming increasingly urgent.

There exist many theoretical frameworks for defining and evaluating the ethical implications of scientific and technological enhancement (Sadler, 2010; Allhoff et al., 2009; Chadwick, 2009; Savulescu, 2006). It will be too ambitious for this essay to discuss the ethical frameworks proposed in the literature. Very shortly, the author's position is that the limit should not be between development and enhancement, since development is, actually, enhancement. Neither should the limit be sought in the difference between therapy and non-therapy, since clothes or automobiles are not therapeutic devices, but, nevertheless, can be legitimately considered as forms of human enhancements. It is the author's contention that every technology, whether material or immaterial (such as language) and all scientific discoveries that heal or improve the condition of human beings should be considered as a form of human enhancement.

In this essay, I argue that enhancement is characterised by a constant and, in many cases, necessary escape from the natural order of things caused by the mediation brought about by science and technology. Quoting from the call of paper of this special issue, human enhancement consists in 'the replacement of the order imposed by nature with the human order'. With technological devices and scientific knowledge human beings can alter the relationship with the natural environment and can modify the nature of their biological bodies. On the one hand, altering the relations with the natural environment and modifying the human body is necessary, since it allows human beings to survive and, on the other, problematic, since it reduces human presence in the world and bring about power relations among living beings and between human beings and the environment. On the contrary, I argue, presence always implies a moral dimension given by the reciprocal structure of the relation.

The purpose of this article is to provide an alternative theoretical framework for explaining and evaluating the ethical implications of technological and scientific forms of human enhancement. While most of the

---

[2] Today, the interaction among genetics, neuroscience, ICTs, and robotic technologies, is making possible even to think of "building" human beings.

discussions on human enhancement focuses on biotechnologies (e.g. Ritalin), this article is concerned with robotics technologies.

The article is organized as follows: in the next section, I shall outline the main features of the theoretical framework proposed and explain the concepts of presence and mediation. In Section 3, I shall apply the framework to robotics technologies and discuss its main effects.

2. The Theoretical Framework – Presence and Mediation

In my opinion, enhancement cannot be discussed without firstly analysing its two main components: on the one hand, the human being and, on the other, science and technology. In other words, without taking into account both human presence – that is, the way of being in the world – and technological mediation - that is, the way in which science and technology mediate human beings' actions and perception of the world, namely presence.[3]

The theoretical framework I propose in this essay is based on the concept of presence, which I understand as fundamentally connected to that of reciprocity, and on the concept of mediation, which I see characterised by two apparently opposite but complementary effects: *extension* and *detachment*. In the next section, I will explain in more details both concepts.

2.1. Presence as reciprocity

Presence is a concept used in many different fields, from philosophy and religion, to business, media, and art. In order to describe what I mean by presence, I will use some insights from the art world, in particular from the field of theatre.[4]

In theatre practice, the word presence can have different meanings. It can refer to the actor's "charisma" (i.e. authenticity or possession), or to the simple fact of "being in the presence of somebody", that is, sharing a space and a certain amount of time with somebody else.[5] However, there is an explication

---

[3] With the word technology I refer also to science, which is the 'underpinning of technology' (ITEA, 2000/2002).

[4] Is it not the theatre the mirror of the world, the metaphor – *par excellence* – of the human life? Calderon de la Barca's 'el gran teatro del mundo' and Shakespeare's 'all the words is a stage' are just two of the most popular statements confirming the strong relationship linking the art of theatre and that of life together.

[5] For a more detailed discussion of theatrical presence see Roger Copeland's already cited study (Copeland, 1999).

of presence which I consider illuminating both for describing theatrical presence as well as for explaining human presence in the world. I refer to the definition given by Roger Copeland who links the phenomenon of presence to the principle of reciprocity: 'presence in the theatre has [...] to do with [...] the way in which the architectural and technological components of the performance space either promote or inhibit a sense of "reciprocity" between actors and spectators' (Copeland, 1990).

Such a conceptualization of theatrical presence shows a fundamental difference with respect to all the explanations mentioned above, since presence is no more considered as a subjective property possessed by either actors (i.e. charisma or authenticity) or spectators (being in the presence of the actors), but it becomes a "relational and mutual experience", which is given by objective circumstances.[6]

Drawing on such a conceptualization of theatrical presence, I propose to understand the more general phenomenon of human presence as *the way in which natural (physical as well as cognitive) and/or artificial (i.e. deriving from culture, including science and technology) factors/conditions either promote or inhibit a sense of reciprocity between two or more people or between a person and the environment* (Salvini, 2006).

Reciprocity is the condition which allows a mutual exchange, immediate or deferred, between a subject and a part of his/her body, between a subject and another subject (i.e. human or animal) and between a subject and the natural environment. For instance, when I touch something or somebody, since I am part of the tangible world, I am also touched by that something or that somebody. Likewise, when I see something or somebody, since I am part of the visible world, I am also "seen" by that something or that somebody.[7] I refer to such a mutual condition as presence.

To conceive of presence as characterised by reciprocity, is to question the explications of presence based on subjective perceptions and unilateral activities.[8] I argue that being in the world is not a univocal act of mind or body,

---

[6] In addition to the architectural and technological components of the theatrical space, I consider determinant for the production of presence in the theatre also the subjective circumstances of both actors and spectators, for instance an actor's bad mood may inhibit the sense of reciprocity with the spectator.

[7] I refer to Maurice Merleau Ponty's notions of "intertwining" and the "flash" (Merleau Ponty 1968[1964]).

[8] Explications of presence based on subjective perception and unilateral activities are widespread in media studies, where presence is understood as the experience of perceiving the environment through either first order or second order mediated perceptions, that is, by the human senses and perceptual processes as well as by human-made technology (Lombard and Ditton, 1997). On the contrary, my contention is that presence is

but is a relational act, which involves both the mind and the body simultaneously in a reciprocal exchange with others and the world: 'our being in the world is far more than just "being". It is a "presence", a "presence" that is relational to the world and to others (Freire, 2001).

Although up until now no study on presence has postulated the relevance of reciprocity in the working and experiencing of presence, I argue that such link can be inferred by taking into account many different disciplines. In philosophy, Maurice Merleau-Ponty (1968 [1964]), Edgard de Vries (1968), Mikhail Bakhtin (1981), Jacques Derrida (1982), Martin Bauber (1987); in cybernetics Norbert Wiener (1954),[9] media studies Jean Baudrillard (1981); in evolutionary and biological studies Humberto Maturana and Francisco Varela (1980); Gregory Bateson (1972); in cognitivist sciences James Gibson (1966), etc.[10] In a study on evolutionary biology, the principle of reciprocity has been described as the mechanism operating in natural selection and explaining the existence of biological systems, such as human society, which, contrary to current assumptions which favour strong and selfish attitudes, are based on altruistic and cooperative interactions (Nowak & Sigmund, 2005). However, as pointed out by Tom Lombardo:

> Not only has reciprocity served as a primary mechanism for the creation of biological and social complexity, but it provides a universal principle upon which human values and ethics are defined. Reciprocity is the foundation of the concepts of justice, equity, and perhaps even human care and kindness (Lombardo, 1987).

Indeed, the condition of presence implies always an ethical dimension – one of responsibility – which arises from the awareness of being in a relation of reciprocity with others and the environment (of course, moral responsibility is also culturally and socially conditioned). As a matter of fact, reciprocity implies

to accept something and to return something). However, to return something

---

a relational experience (Salvini 2006). Indeed, there are artists in the field of "telepresence art", such as Roy Ascott, Paul Sermon, and Eduardo Kac, whose intention is to replicate presence at a distance not by providing "users" with immersive or interactive experiences, but by recreating reciprocal relations (Salvini, 2005).

[9] In particular, I found interesting the debate between "vitalist" and "reductionists" and the prominence given by both schools of thought to the feed-back loop, or self-regulation capability, that is, to the ability to adapt to the environment, which is given by reciprocal connections (Cordeschi, 1998).

[10] This is not an exhaustive list.

implies to possess something to give back but also, and most importantly, the possibility to have access to the receiver. As pointed out by Jean Buadrillard:

> The totality of the existing architecture of the media founds itself on this pattern definition: they are what always prevent response, making all processes of exchange impossible [...]. This is the real abstraction of the media. And the system of social control and power is rooted in it (Baudrillard, 1981: 170).

According to Baudrillard, the lack of responsibility corresponds to an impossibility to reciprocate:

> Power belongs to the one who can give and cannot be repaid. To give, and to do it in such a way that one is unable to repay, is to disrupt the exchange to your profit and to institute a monopoly' (Ibid., 170).

Therefore, besides an ontological condition, presence implies also an inherently political condition, characterised by an equal balance of powers and by a moral attitude among the subjects of the relation.

As I shall point out in the next sub-section, technologies affect the condition of reciprocity typical of presence and can alter the power relations.

In order to understand how science and technology destabilize presence and bring about power relations, it is necessary to focus on their effects of mediation.

2.2. Technology as Mediation

'Technology extends our abilities to change the world: to cut, shape or put together materials; to move things to one place to another; to reach further with our hands, voices and senses' (AAAS, 1993). Technology and science can be analysed by taking into account their double function: as instruments to achieve a goal, but also as forms of mediation of human actions and perceptions of the world. As a matter of fact, by mediating human actions, science and technology may also affect the way we relate and experience the world.

According to the framework proposed in this essay, the act of mediation performed by technology neither is considered for its instrumental function, nor for its epistemological or hermeneutic effects. On the contrary, the focus is on human presence, that is, on how technology affects the structure of the relations of reciprocity that exist among human beings, the human body, and the natural environment. My contention is that changes in the structure of reciprocity may alter or determine power relations, moral disengagement,

abstraction and loss of values. Before turning to these issues, I need to explain in more details the act of mediation.

There seem to me to be two important aspects, apparently opposite but complementary, characterising any form of mediation, which I define the "double logic of mediation".[11] First, there is the *act of extension*, which consists in extending a physical or cognitive capability (or creating a new one!). Second, there is *detachment*, which consists in creating a distance between a human being and his/her own body, others or the outside world.

In his well-known analyses of media, Marshall McLhuan pointed out that 'the content of a medium is always another medium' (McLhuan, 1994 [1964]). According to the act of extension, for instance, a car is an extension of a carriage, which on its turn is an extension of a horse, which is ultimately an extension of our legs; likewise, the house can be considered as an extension of a dress, and a dress an extension of our skin. However, as I have mentioned before, besides the act of extension, a medium is also and always characterised by the act of separation or detachment. Indeed, by definition, any medium is characterised by the fact of staying in between something.[12] Therefore, by extending human beings' capabilities and perceptions, technology is also performing an act of separation (or detachment) between a person's action and his/her body.

The live screen can be considered as the archetype of the "double logic of mediation", as well as of its main negative effects, namely empowerment, abstraction and moral disengagement. The camera and the screen can be considered as an extension of the eye, but while bringing "things" closer and "at hands", at the same time, they keep them apart. The screen can be a window and at the same time a shield. It allows someone to see, but at the same time it screens out the visible. According to Kevin Robins, the screen is a space of visibility and invisibility:

> [t]he nature and functioning of the screen are crucial. The screen has allowed us to witness the world's events while, at the same time, protecting us – keeping us separate and insulated – from the reality of the events we are seeing. [...] The force of the screen works to make moral response more difficult

---

[11] The "double logic of remediation" is a concept originally devised by Jay David Bolter and Richard Grusin in their study *Remediation: understanding new media* (Bolter and Grusin, 1999). According to their theory, the double logic refers to the desire for immediacy and, at the same time, for hypermediacy characterising new media. I have adapted their concept to my arguments.

[12] A "medium" is defined as: 'a state that is intermediate between extremes; a middle position' (WordNet, 2014).

(Robins, 1996).

With regards human presence, the double logic of mediation can either promote or impede the structure of reciprocal relations. To promote reciprocity means to replicate, multiply (also by adding new bi-directional channels) or extend the existing channels of reciprocity given by our biological body, which I consider the natural enabler and carrier of presence. For instance, while extending visual and audio capabilities a telephone or a teleconference system allow to bridge the physical distance between two or more interlocutors. However, to conceive of both of them as fully based on a reciprocal structure is to overlook the fact that the world of the tangible, that is, the possibility to touch one-another typical of an in presence situation, is completely missing. Therefore, whatever the motivations (e.g. technological limitation, functional restriction, or deliberate choice),[13] technological mediation always disrupts presence.

3. Discussion: Robotics, Presence, Mediation and Human enhancement

In the remaining part of this section, I shall apply the theoretical framework based on presence and mediation to tele-operation systems (aka tele-robotics) and discuss its relevance for the ethical evaluation of human enhancement. For the reader's convenience, I shall start by briefly describing the main features of a tele-operation system.

While it is easy to recognise a robot, to say what a robot is can be very difficult.[14] I consider programmability,[15] that is, the possibility to inscribe a certain behaviour or a set of future actions in an artefact, what defines and distinguishes a robot from another object.[16] Very simply, a robot can be

---

[13] According to Langdon Winner 'in many instances, to say that some technologies are inherently political is to say that certain widely accepted reasons of practical necessity [...] have tended to eclipse other sorts of moral and political reasoning (Winner, 1980).

[14] In Joseph Engelberger's words: 'I can't define a robot, but I know one when I see one' (Engelberger, 1989).

[15] It is worth noting that according to this definition, the possibility to re-program human genetic code by means of genetic engineering (as advocated by many trans-humanists) would be more appropriate to define a human as a robot, rather than the replacement of human organs with artificial (robotic) ones.

[16] According to the International Standard Organisation (ISO), a robot can be defined as an 'actuated mechanism programmable in two or more axes with a degree of autonomy, moving within its environment, to perform intended tasks' (ISO 8373, 2012).

programmed to be autonomous or tele-operated.[17] The level of autonomy is directly related to the concepts of presence and mediation, since it concerns the degree of human involvement in the tasks or actions carried out by the robot. As a matter of fact, in robotics engineering, autonomy is defined as the robot ability 'to operate in the real-world environment without any form of external control, once the machine is activated and at least in some areas of operation, for extended periods of time' (Lin, Keith, & Beckey, 2011).[18]

A tele-operation system consists of four main components: the human operator, the interface at the local site (i.e. the master) – usually a contact device which allows the user to send commands and receive sensor signals – the communication link (in tethered tele-operation systems, information is exchanged by cable, whereas in untethered systems by wireless connections), and, finally, the robotic artefact at the remote location (the slave) – provided with actuators, controllers, a power system, and various kinds of sensors, which vary depending on the remote task (Goradia & Elhaj, 2005). Nowadays, notwithstanding the great advances in computation and perception, tele-operation systems are still used in many fields of application, among the most popular are: search and rescue operations, surgery, space and underwater explorations, and warfare.

Robotic prostheses for upper or lower limbs and exoskeletons are a special type of tele-operation system. With respect to conventional tele-operation systems, they have a different ontology. As a matter of fact, during operation the robotic device is always in contact with the human body.

A tele-robotic system extends the human operator's range of actions (by extension I mean that it can enhance motor and perceptual capabilities) in an otherwise inaccessible environment (due to distance, scale, danger, etc.). Therefore, the human being's actions are no more bounded to the *hic et nunc* of the physical body, but they can take place *there and now*. In the case of a robotic hand prostheses or an exoskeleton there is no physical distance between the human operator and the robot, since the robot is in contact with the human body. As far as the control interface is concerned, robotic prostheses and exoskeletons are usually controlled by decoding muscular (electromyographic) or neural (from peripheral or central nervous system)

---

[17] There are also intermediate levels of autonomy, such as semi or shared autonomy, which, for simplicity, I will not take into account.

[18] Elsewhere, I have argued that an autonomous robot is the utmost kind of human enhancement, since it completely detaches a human being from the presence in the world (Salvini, 2012).

signals via dedicated invasive or non invasive electrodes. With respect to a joystick or another contact interface, muscular and neural interfaces allow a more natural and intuitive control of the robot and increase the sense of ownership between the device and the human body.

For instance, a tele-operated robot designed for the inspection of a nuclear power plant allows a human being to accomplish tasks from a safe location (i.e. the control station) and even from a different time-zone. An exoskeleton designed to assist people in walking activities can support or increase the physical capabilities of a person, while a prosthetic hand may restore some of the amputee's lost capabilities.

These and many other applications are usually considered as the typical ways in which robotics brings about forms of human enhancement. As a matter of fact, they can be understood as attempts to overcome the current limitation of human body through natural or artificial means.

However, in my opinion, there is a more profound sense in which robots can contribute to the enhancement of human beings, which I will illustrate by applying the theoretical framework proposed in this article. As pointed out in Section 1, it is necessary to consider how technological mediation affects the structure of reciprocal relations. In other words, how technological mediation changes human presence, that is, the way of being in the world.

In applying the framework, I will point out the relevance of the issue of power, which is usually under evaluated in the discourses surrounding human enhancement, which are mainly focused on issues of freedom and autonomy, health and safety, fairness and equity, societal disruption and human dignity (Lin & Allhoff, 2008).

Drawing on the double logic of mediation, in a tele-operation system, on the one hand, the robot extends the human operator's action in the remote environment. By extension, I mean not only overcoming spatial distance, but also going beyond human capabilities, such as to increase human strength or sensory perception (e.g. night vision, echolocation, etc.) or to overcome physical limitations, as in the case of an amputee. However, on the other hand, the robot keeps the human being at distance from the world, by detaching him/her from what is happening in the remote environment (e.g. from the effects of his/her action). In other words, the system disentangles the human user from the world of the visible and that of the tangible, by turning the structure of the relation of contact and vision from one of reciprocity into one of unilaterality.

For instance, the handling of hazardous material via a tele-robotic system or the possibility to lift heavy loads by means of a robotic exoskeleton are all based on the inhibition of the channels of reciprocity usually active in a situation of presence. The technological mediation prevents the person wearing the exoskeleton to feel the weight of a load – and being squashed by it – or, in a tele-operation system, to undergo the effects of the radiations present in the working environment. Finally, in the case of a robotic prosthesis, the technological mediation allows the subject to restore a lost functionality (e.g. grasping or manipulating objects), but, at the same time, it reiterates the distance between the amputee's body and the external world. For instance, the lack of sensory feed-back or the design of sensory feed-back system not based on biological models can determine distance.

Parenthetically, the quest for designing reciprocity into artefacts is illustrated by current researches in developing sensory feed-back in robotic prostheses for delivering real-time sensory information to upper limbs amputees (Raspopovic et al., 2014). The link between a person's intentions and his/her actions is given by an artificial system capable of decoding and coding the signals coming from the human. The feed-back system is based on a decoding subsystem for detecting the user's intentions via efferent nerves and a delivery sub-system which uses afferent nerves.

At present, the main obstacles to the development of a full sensory feed-back system are due to technological and scientific limitations. However, it is very likely that in the future, such a limitation could be exploited and turned into a form of empowerment for the person wearing the device. As a matter of fact, a robotic hand prosthesis could be designed to allow a person to accomplish tasks beyond the biological properties of a fleshy hand in order to, for instance, resist extreme hot or cold temperatures, pain and fatigue. Therefore, once it will be possible to design an almost natural prosthesis, it is questionable whether there will be the need of laws for regulating the types of threshold, i.e. biological or artificial, to be implemented in the robotic prostheses.

In conclusion, if the possibilities offered by real-time communication technologies (i.e. the digital screen), have disrupted the way of seeing the world and altered the moral response of human beings, mechatronics and A.I. (i.e. robotics) have disrupted the way of being in touch the world. In almost all the tele-robotic systems discussed above, the sense of "touch", the most intimate among human senses, is no more characterized by a relation of reciprocity (i.e. to touch always implies being touched), but it is artificially

configured in a relation of distribution, in which the "afferent" property has been "severed" from the "efferent" property. In handling a thing by means of a robotic avatar, an exoskeleton or a robotic prosthesis, the technological mediation stays in between our own bodies and the object of the action. Moreover, the mediation may be designed so as to prevent any form of response. Therefore, the reciprocity of human touch is lost in the mechanically reproduced forms of touch (i.e. tele-touch).

Far from being a neutral replacement of bodily touch, "tele-touch" allows human beings to advance on the ladder of progress by keeping the world at a distance. It has already been pointed out that search and rescue operations, surgery, space and underwater explorations and many more tasks are now possible or have been improved thanks to tele-robotics systems. However, tele-touch can also become a dangerous instrument of power, since, as pointed out by Francis Baudrillard, it prevents the possibility to respond back and make the process of exchange impossible. Because they exploit such a possibility, warfare technologies can be considered as a case in point.

The loss of reciprocal relations generates a situation of power, which may be determinant for the accomplishment of a task, but it is detrimental for the sense of presence. Furthermore, as pointed out in section 1.1, presence always implies a moral dimension, which is the result of being in a reciprocal relation with the other or the environment. However, by staying in between, technological mediations may weaken moral response. According to Kevin Robins, 'our technologies keep the world at a distance. They provide the means to insulate ourselves from the disturbing immediacy of the world of contact' (Robins, 1996).

Finally, technological mediation allows to modulate the channels of reciprocities according to new thresholds and filters, which are not based on the default biological settings (i.e. those of the human body), but are given by the properties of the technological components. The result is a new artificial capability that may alter the system of human and physical values, since it allows a person to do things without experiencing the corresponding consequences.

4. Conclusion

To sum up, in this essay I have argued that human enhancement, which I consider the result of technological and scientific progress, is part of

humankind's nature. By protecting the frail condition of the human body and increasing its limited capabilities, science and technology keep the world at a distance, thus reducing human presence. To reduce human presence means to change the structure of the relations of reciprocity that exist among and within living organisms and between them and the natural environment. The result is a relation in which only one side is able to act and feel while responses from the other side are negated. The new structure may alter the power relations and bring about moral disengagement, abstraction, solipsism, in one word absence from our own selves, our body, other living beings and the natural environment.

The framework proposed for making sense of human enhancement and for evaluating its ethical consequences is based on two concepts: presence, which I have explicated as consisting in a network of reciprocal relations determined by natural as well as artificial conditions, and mediation, which consists of two complementary, but opposite aspects: on the one hand extension, which tends towards unification and on the other detachment which tends towards separation.

According to the workings of the double logic of mediation, even in the case of a medium like the telephone, in which the dialogic structure of a face to face conversation is replicated, there is, nevertheless, always a loss of reciprocal relations, that is, the impossibility of achieving a full presence. In fact, while bringing together the two voices, the telephone prevents any form of contact or visibility.

Among the many plausible objections to the theoretical framework I proposed in this essay is that it replicates a logocentric pattern. The natural body and the immediate (i.e. non-mediated) presence being the normative starting point. However, according to my argument, the starting point should not be confused with the normal functioning of human beings. The concept of presence I have proposed in this essay is based on a reciprocal relation. Therefore, it is the relation the central feature of the framework rather than the natural body. Nevertheless, it seems to me impossible to avoid considering the human body as the enabler and carrier of human presence.

Furthermore, one could argue that it is questionable whether the framework would be applicable also to non-technological forms of human enhancement, such as Ritalin. My contention is that all forms of progress, whether technologically or scientifically enabled, can be considered as mediations and therefore are under the spell of the "double logic", which

extends and at the same time reduces human presence. In the case of chemical forms of enhancement, such as Ritalin, the drug provides a subject with more energy and memory and allows him/her to perform better. However, the drug affects also the subject's ability to respond to the physical and psychological effects triggered by high cognitive or physical performances. For instance, Ritalin prevents the subject from experiencing the sense of fatigue, which in "normal physical conditions" is connected to the awareness that high performances need training, perseverance and determination. The structure of the relations of reciprocity within the body has changed: the capabilities of the human body are increased (extension), but at the same time the physiology of the human body has been inhibited (detachment). Likewise, a robotic exoskeleton can allow someone to raise heavy loads, but the user, while using the robot, will lose the sense of heaviness of things, since the capabilities of the robot are not set in accordance with the physical and psychological properties of the human body but, on the contrary, the thresholds of the possible and impossible will be given by the properties of the machine.

In more general terms, a chemical drug or a robotic exoskeleton, by diminishing physiological responses, such as the feeling of fatigue or the sense of weight, may produce serious consequences on the power relations, the health of an organism, as well as change the value of things. As pointed out by Francis Fukuyama, human nature 'is fundamental to our notions of justice, morality and the good life' (Fukuyama, 2002).

Robotics technology has allowed human beings to accomplish incredible things and to advance human knowledge. However, the double logic of mediation subtends all kinds of science and technology. It will be our choice to decide whether technological mediation should bring humankind towards presence or towards absence. It will be a matter of education, design, and politics.

REFERENCES

AAAS (1993). *Benchmarks for Science Literacy*. American Association for the Advancement of Science.

Bahktin, M. (1981). The Dialogic Imagination: Four essays. In Holquist, M. (Ed.). *Dialogism. Bakhtin and his World*. London: Routledgde.

Bateson, G. (1972). *Steps to An Ecology of Mind: Collected essays in anthropology, psychiatry, evolution and epistemology.* Aylesbury: Intertext.

Baudrillard, J. (1981). *Requiem for the Media. For a Critique of the Political Economy of the Sign.* St. Louis: Telos Press.

Buber, M. (1959). *I and Thou.* Trans. by Ronald Gregor Smith, 2nd edition, Edinburgh: Clarke.

Bolter, R., & Grusin, D. (1999). *Remediation. Understanding New Media.* Cambridge, MA: The MIT Press.

Chadwick, R. (2009). Therapy, enhancement and improvement. In Gordijn, B., & Chadwick, R. (eds.) *Medical Enhancement and Posthumanity* Springer, 25-37.

Copeland, R. (1990). The presence of mediation. *The Drama Review, 34*(3).

Cordeschi, R. (1998). *La scoperta dell'artificiale. Psicologia, filosofia e machine intorno alla cibernetica.* Milano: DUNOD

Derrida, J. (1978). *Writing and Difference.* Trans. by Alan Bass, London: Routledge.

Engelberger, J. (1989). *Robotics in Service.* Cambridge, MA: The MIT Press.

Freire, P. (2001). *Pedagogy of Freedom: Ethics, Democracy, and Civic Courage.* Rowman and Littlefield Publisher Ltd.

Fukuyama, F. (2002). *Our Posthuman Future: Consequences of the Biotechnology Revolution.* Picador.

Galimberti, U. (1999). *Psiche e tecne. L'uomo nell'età della tecnica.* Milano: Feltrinelli.

Gehlen, A. (1988 [1940]). *Man, His Nature and Place in the World.* New York: Columbia University Press.

Gibson, J.J. (1966). *The Senses Considered as Perceptual Systems.* Boston: Houghton Mifflin Company.

Goradia, A., Elhaj, N., & Xi, I.H. (2005). Internet Based Robots: Applications, Impacts, Challenges, and Future Directions. *IEEE Workshop on Advanced Robotics and its Social Impacts.*

ISO8373 (2012). Robots and robotic devices – Vocabulary. International Standard Organisation.

ITEA (2000/2002). International Technology Education Association.

Lin, P. & Allhoff, F. (2008). Untangling the debate: The Ethics of Human Enhancement. *Nanoethics*, *2*(3), 251-264.

Lin, P., Abney, K., & Bekey, G., (2011). Robot ethics: Mapping the issues for a mechanized world. *Artificial Intelligence*, *175*, 5-6.

Lombard T., & Ditton, M. (1997). At the heart of it all: the concept of presence. *Journal of Computer Mediated Communication*, *3*(2).

Lombardo, T.J. (1987). *The reciprocity of Perceiver and Environment. The evolution of James J. Gibson's ecological psychology*. Hillsdale, NJ: L. Erlbaum Associates.

Maturana, H., & Varela F.J. (1980). *Autopoiesis and cognition: the realization of the living*: Boston: D. Reidel Pub. Co.

Merleau-Ponty, M. (1968[1964]). The Intertwining-The chiasm. In Lefort, C. (ed.) *The Visible and the Invisible.* Trans. Lingis, A., Evanston, Ill.: Northwestern University Press.

McLhuan, M. (1994 [1964]). *Understanding Media: The Extensions of Man.* Cambridge Massachusetts: The MIT Press.

Menuz, V., Hurlimann, T., & Godard, B. (2011). Is Human Enhancement also a Personal Matter? *Science and engineering ethics*.

Nowak, M.A., & Sigmund, K. (2005). Evolution of Indirect Reciprocity. *Nature*, *437*.

Peursen Van, C.A. (1968). Notes for a philosophy of reciprocity. In de Vries, E. (ed.) *Essays on Reciprocity.* Institute of social studies, Paris: Mouton the Hague.

Robins, K. (1996). *Into the Image: Culture and Politics in the Field of Vision*, London and New York: Routledge.

Sadler, J.Z. (2010). Dignity, arete, and hubris in the transhumanist debate. *American Journal of Bioethics*, *10*(7).

Salvini, P. (2005). *An analysis of Roy Ascott, Eduardo Kac, Paul Sermon, Norma White and Doug Back's telepresence art: a disorganised space of potential reciprocity.* Dissertation in part fulfilment of the requirements for the degree of MA by research (M.Res) in Theatre Studies, Lancaster University, September (unpublished).

Salvini, P. (2006). Presence: A Network of Reciprocal Relations. *PRESENCE 2006 The 9th Annual International Workshop on Presence*, Ohio, Cleveland, August 24-26.

Salvini, P. (2012). Presence, Reciprocity and Robotic Mediations: The Case of Autonomous Social Robots. *International Journal of Technoethics*, *3*(2), 9-16.

Savulescu, J. (2006). Justice, fairness, and enhancement. *Annals of the New York Academy of Sciences*, *1093*, 321-38.

Stearn, G.E. (1968) (Ed.). *McLuhan Hot & Cool*. Harmondsworth: Penguin.

Virilio, P. (1997 [1995]). *Open Sky*, trans. by Julie Rose, London and New York: Verso.

Virilio, P. (1995). Absolute velocity. *MediaMente* (interview, Paris - European IT Forum, 05 September) [Online] Available: http://www.mediamente.rai.it-/mmold/english/bibliote/intervis/v/virilio.htm.

Vries de, E. (1968a). Explorations into reciprocity. In de Vries, E. (ed.) *Essays on Reciprocity*, Institute of social studies, Paris: Mouton the Hague

Vries de, E. (ed.) (1968). *Essays on Reciprocity*. Institute of social studies, Paris: Mouton the Hague.

Wiener, N. (1954 [1950]). *The Human use of Human Beings – Cybernetics and society*. Da Capo Press.

Winner, L. (1980). Do Artifacts Have Politics? *Daedalus*, *109* (1).

WordNet (2010). Princeton University. [Online] Available: http://wordnet.prince-ton.edu.

Book Review

Human being @ Risk. Enhancement, Technology, and the Evaluation of Vulnerability Transformations

by Mark Coeckelbergh, Springer, 2013

Jacopo Branchesi[‡]
j.branchesi@sssup.it

How do persons and societies cope with risk and vulnerability? This is the fascinating philosophical question to which Coeckelbergh's book, *Human beings @ risk* intends to respond. Human beings are at risk since they are in the world: vulnerability, Coeckelbergh claims, is condition of being in the world, in other words it is something inherent to human condition. At the same time, human beings are always already engaged in the struggle against risk and vulnerability in order to minimize or radically eradicate them from their life: acting to reduce vulnerability is an instinct implanted in human beings by nature, which drives human behaviours in social environment like the Hobbesian instinct of self-preservation.

The issue seems to be clear and linear enough: both vulnerability and the human struggle against vulnerability represent naturally and historically existential conditions, they are strictly related and the latter is a consequence of the former one. Nevertheless, the relationship between human being and vulnerability is more complex than what may appear at first sight, since the human struggle against vulnerability entails a reflection on the concept of human enhancement and a careful evaluation of related ethical problems: is it ethically acceptable to overcome the human limitations by the means of technology in order to remove or radically minimize vulnerability? "And can we become invulnerable? Are there limitations to 'making the human'?" (Coeckelbergh, 2013, p. 10).

The ongoing debate on human enhancement and its assumption is highly polarised between who defends the human nature (conservatives like Habermas, Fukuyama, McKibben, Elliot), on one hand, and who wants to

---

[‡] Scuola Superiore Sant'Anna, Pisa, Italy.

change human nature, to improve human capabilities moving towards the post-human age (transhumanists like Bostrom, Ord, Stock, Kurzweil), on the other hand.

In his essay, *Human Being @ Risk*, Coeckelbergh does not intend to argue for a position in this debate, but he tries to overcome the static nature of the positions and to mitigate the polarisation in the discussion, by developing an anthropology of vulnerability, which includes a normative dimension of anthropology: thus, the question "is not only what we are as humans (descriptive), but also what we should be (normative)" (Coeckelbergh, 2013, p. 21).

According to the author, technologies hence human being and transform human vulnerability – they already do it – but do not eradicate it. Human vulnerability, he claims, is ineradicable and we may just transform it by technologies; therefore, he invites us to think about "what kind of vulnerability we want" (Coeckelbergh, 2013, p. 10) in relation to what we should be or want to be.

Within the philosophical perspective elaborated by the author, the anthropological question becomes a political question and lately a question related to democratic praxis in our contemporary societies.

Coeckelbergh's essay is focused on the philosophical and anthropological analysis of the critical relationship between vulnerability and technology: the aim is help us to cope with the existential condition of vulnerability, without sacrifice a dynamic idea of human being as perfectible being who tries to go beyond his limits.

Now, let us observe the book in outline: it is articulated in three strictly connected parts, through which the author develops his normative anthropology of vulnerability.

The first part discusses the descriptive anthropology of vulnerability. After an overview of the debate about human enhancement, in which the author briefly retraces the main arguments of both positions transhumanists and conservatives, he proposes a dynamic interpretation of human being, which allow him to get out of the conservative position without falling in the radical transhumanist one: human nature has always evolved, human being has always been changed by technologies, nevertheless it does not mean that any kind of human transformation and improvement is possible (Chap. 2).

Within this anthropological framework, Coeckelbergh conceives vulnerability as existential condition (Chap. 3), interpreting it by both the

Heidegger existential categories and Jackson empirical anthropology. The idea of vulnerability, as existential condition, makes rise the need of a culture of vulnerability (Chap. 4), which is discussed by the author through an analysis of the representations and practical experiences of vulnerability.

In the second part, Coeckelbergh moves from the descriptive anthropology to the elaboration of the normative anthropology of vulnerability. He takes in consideration firstly the ethics of vulnerability (Chap. 5), pointing out that there is no opposition between ethics and technology, since both represent different ways of our attempt to understand and figure out the deepest meaning of vulnerability.

Following this line of thought, the author suggests to observe how human vulnerability has already been transformed by technology; to examine in which sense we are at risk today in the light of the past and, finally, just at that point we are able to evaluate the material and ethical meaning of these transformation operated by technology.

At this point, Coeckelbergh examines the meaning and implications of the ethics of human enhancement (Chap. 6), focusing on the transformation of our values as a consequence of the new possibilities opened by technologies. Then, he turns his attention to the aim of transhumanist project: by exploring possible sceneries of a post-human development, he points out that hypothetical post-humans would be still vulnerable; and he concludes that if the aim of transhumanist project is to eradicate vulnerability, then it is doomed to fail, since vulnerability can just be transformed and never be eradicated at all.

Even looking at today technologies, like Internet or some technological applications in medical field, we have to recognize that while they reduce past forms of human vulnerability, they create new and different kinds of vulnerability (Chap. 7).

Therefore – the author concludes – given that technology already change human nature and taking in consideration that vulnerability can not be eradicated, then the crucial questions are the following: which changes and improvement of human nature do we want and do we consider sustainable? And, consequently, what kind of vulnerability do we prefer? To what extent do we still recognize ourselves as humans? According to Coeckelbergh, these questions are not merely private, but they are public affair: issues regarding human nature like human enhancement, vulnerability and the use of technology on human being should be addressed to society and discussed within the public space.

In this manner the author introduces the notion of the politics of vulnerability: what does it intend? He identifies different meanings of this notion, from the politics of human rights to the new forms that politics should assume in order to face the new kinds of vulnerability created by technology (Chap. 8).

Remaining on the ground of politics, then Coeckelbergh focuses on the relationship between vulnerability and the three key concepts of political philosophy: freedom, justice and democracy. Exploring the moral and political implications which rise from these relations, he deals with Hobbes thought, especially the man antisocial instincts and the conception of risk, the Rawls two principles of justice, and Arendt theories on democracy and political freedom. What it emerges is that new forms of vulnerability continuously rise and they generate new problems to those concepts of political philosophy which have moulded the modern and contemporary public space.

In the last chapter (Chap. 9) of the second part the author elaborates the idea that coping with vulnerability is an art, and as an art it requests skills: how to acquire these skills? He claims that such skills can be acquired by adopting a relational-ecological conception of existential vulnerability, which takes a shape in the book.

In the third part, Coeckelbergh reunites the arguments he has articulated during the essay and concludes that we may cope with vulnerability firstly by reflecting on the descriptive question about the history and future of vulnerability transformations, and then by reflecting on the normative question about which vulnerability we want, that finally means reflecting on the question: what kind of society do we want? On the contrary, according to the author, thinking about the dream (or nightmare) of invulnerability like an end of the human enhancement has no sense at all, since the book shows that vulnerability is embedded in our relation with the world. Paraphrasing the Achilles story told us by the ancient Greek myth, Coeckelbergh at the end claims: "we are – and we will remain – at once the heel and the arrow" (Coeckelbergh 2013, p. 204).

Finally, we may say that the strengths of the essay are: on one hand, the shift from the static notion of human nature to that of human being as a dynamic being; on the other hand, the idea of vulnerability as relational category inherent to our existence in the world. These two arguments allow the essay to overcome the static debate on human enhancement and to open a normative perspective that makes possible an inquiry on the relationship between

vulnerability and technology transformations on the ground of philosophical anthropology, as well as on those of ethics and politics.

REFERENCES

Beck, U. (1992). *Risk society: Towards a new modernity.* Translated by M. Ritter. London: Sage.

Coeckelbergh, M. (2011). Human development or human enhancement? A methodological reflection on capabilities and the evaluation of information technologies. *Ethics and Information Technology*, *13*(2): 81–92.

Habermas, J. (2003). *The Future of Human Nature*. Cambridge: Polity Press.

Marx, K. (1867). *Capital*, vol. I. Harmondsworth: Penguin.

Nussbaum, M., & Sen, A. (1993). *The Quality of Life*. Oxford: Clarendon.

Sartre, J.-P. (1948). *Existentialism and Humanism*. Trans. Mairet, P. London: Methuen & Co.

Book Review
Unfit for the Future

by Persson & Savulescu, New York: Oxford University Press, 2012

Laura Crompton [†]
laura.crompton@campus.lmu.de

In the book *Unfit for the Future* Persson and Savulescu portray the problems
and challenges humanity will have to cope with in the near future. Problems
technological progress and demographic growth have evoked can't be solved
through common moral psychology, which was assimilated to small, non-
technological societies many many years ago. There is a need of moral
enhancement for humanity to be able to cope with present problems.

It is Persson and Savulsecus opinion that humanity is 'ill-equipped' (p. 12)
through the so called 'common-sense morality'. Moral attitudes of various
societies all over the world can be brought to one common denominator, which
the authors call 'common-sense morality'. This 'common-sense morality' is not
capable of giving us the moral psychology to cope with the problems modern
societies have to face. Further in the book, Persson and Savulescu illustrate the
components of this 'common-sense morality'. For example, it is said that we
care more about what happens to the people who are close to us, than the ones
further away. Also it is rather difficult for us to sympathize and emphasize with
a larger group of people. Furthermore our causal contribution towards an
action is proportional to the responsibility we carry for this action. Another
point illustrating the 'common-sense morality' is that altruism, which exists as
long as the members of a community are able to observe each other constantly.
All these components forming this 'common-sense morality' can't be
transferred to modern societies as they consist of millions of citizens in which
anti-social individuals are bound to strive against the stream. In my opinion the
theory of a 'common-sense morality' is very strong and needs to be looked at
more precisely.

[†] Ludwigs-Maximilians-Universität München, Germany.

It is indeed true that the history of all societies goes back to one common point from which on these developed in various ways. But it's these various ways out of which norms, morality and the ethics of a culture sprouted and grew. We learn and adapt certain moral attitudes through the circumstances we grow up in, presented throughout our parents and the environment. This is the morality that influences our evaluation and the actual process of certain actions. Each culture has its own way of handling moral questions, dilemmas and problems. It may be true that moral attitudes at the very beginning of societies were strongly related and showed parallels in between each other, which explains why even now some features might be very similar if not even the same. But as I already mentioned, societies very quickly developed their own cultures and traditions influencing moral attitudes within each individual society. Taking all this into consideration, it seems reasonable that moral attitudes can't be taken under one common denominator, as these sprout of the societies' individual cultures and environments. 'Common-sense morality' can't exist, as moral attitudes are individually adjusted to separate societies. Another point, which needs to be taken into consideration is the fact that we have now have the possibility of sharing and communicating our moral attitude through the connected world we live in, made possible through globalisation. Through social networks such as Facebook or video platforms such as YouTube, we have the possibility of experiencing different moral attitudes from other cultures. In this sense it again is questionable whether there is such a thing as 'common-sense morality'. Or are moral attitudes starting to adjust one to another through our overly connected globalized world?

Further in the book, the authors outline the problems modern societies have to face in the future and how 'ill-equipped' humanity is through 'common-sense morality'. The first topic the authors assign to, are the problems which emerge through liberal democracies. Liberal democracies are defined through the fact that all citizens have the same rights and liberties. Therefore a liberal state has a market economy, freedom of speech and press and freedom of religion. The authors mainly concentrate on the threats and challenges these liberal democracies have to face. In general it's easier to harm an individual, a group or a functioning system than to heal or repair it. As our standards of living are now as high as they have never been before, it will be harder to increase the quality of our living, rather than do harm and therefore decrease it. Technology is an enhancement to humanity, making it easier to threaten or do damage. Consequently, liberal democracies display a great target for such

threats. Analogously to the growth and spread of technological and scientific knowledge, is the information and power that goes with it. As liberal democracies have vastly grown over the years, it gets more difficult to pay surveillance to this information and power, which displays a great danger. The main threats Persson and Savulescu concentrate on are weapons of mass destruction and the anthropogenic climate change. To resolve danger, the security within liberal democracies would have to be strengthened, but through this, one of the basic characteristics of liberal democracies is lost, namely the right for privacy.

What needs to be taken into consideration here is that technology also illustrates a great possibility to humanity. The authors say that alongside the technical revolution came great moral change (p. 117). "when people undergo great moral development in the course of their lives, their moral competence will largely die with them" (p. 118), is what the authors say concerning moral change. Displaying this thesis to the example of the computer wave in the 90ies, this would imply that a part of our morality died. Nowadays thousands of people all around the world have computers. It is obviously true that through these, new possibilities and dimensions were born, there was and still is change within moral attitudes. But proposing that our 'moral competence' has 'largely died' through this is a very strong, and in my opinion wrong thesis. It is indeed true that our societies through technology had to overcome great change within the range of morality. But societies are under permanent change and have to face moral challenges. One could actually say that moral change comes alongside with the developing of a society. The change of woman's rights and equality of man and woman constitutes a good example. Over hundreds of years it was commonly spread that female stood way underneath the male. Since 1850,[1] through many fights and demonstrations, societies experienced great moral change and development. In many parts of the world female are no longer less valuable than males, they have reached the same point of rights and liberties. This was a great change for society, but would one say that through this a large part of our moral competence died? Definitely not, it seems that through this revolutionary change our moral competence grew rather than decreased. In this sense, it is my opinion that great moral development doesn't go hand in hand with our moral competence dying. Facing the moral changes

---

[1] Woman's rights, 2011: http://www.uni-bielefeld.de/gendertexte/chronik.html (Accessed on 08 Januar 2014)

and challenges coming alongside technological progress, one needs to see it as an enhancement to humanity, proposing changes and new possibilities. Indeed we need to keep in mind that these changes could bring harm to our life, which is why we always need to question and be attentive towards moral changes coming through revolutionary innovations.

To strengthen their argument that the technological revolution poses a threat for liberal democracies, further in the text, the authors suggest dropping 'common-sense morality', as well as the act-omission-doctrine. According to this doctrine, one is for example not a murderer if he doesn't commit the actual act of killing somebody. If the person dies because of circumstances in which one doesn't intervene through an act, we omit helping or saving the other persons' life, but are according to the act-omission doctrine not a murderer (see Howard-Synder, 2011, for a summary article). It is the authors' point of view that both of these don't work in a globalized world as ours. Through the technological progress, for example, our actions get more powerful. Analogously to the increase of the power of our actions, the range of what happens when we omit to use these powers grows. Concluding, our moral responsibility grows through the power our actions have. The authors suppose that we are just as responsible for occurred harm if we caused it, as if we had known about it, but omitted preventing it. Persson and Savulescu say that through the fact that societies nowadays are so big and individuals act amongst each other, our view and feeling of responsibility is strongly diluted. Thus, we for example don't feel responsible for the misery in developing countries, primarily because there are many agents involved who could, as well as we set action against it. Again, the 'common-sense morality' is proofed of being too weak to cape with the present problems of modern societies. In the following chapter Persson and Savulescu outline the 'Tragedy of the Commons', the overly big societies outsourcing our resources leading to anthropocentrically caused environment and climate changes. The economic growth and increasing population numbers within the liberal democracies are a great contribution to the present condition of our planet. As the number of people involved in the 'Tragedy of the Commons' is big, the sacrifice of relinquishing something gets more distant for each individual. Within societies as big as they are nowadays, there are many free-riders and anti-social beings striving against the stream and thus causing distrust amongst each other. We are aware of their existence and know they would not sacrifice anything to them meaningful, through which it gets harder for us to do so. In order to avoid the 'Tragedy of

the Commons' many parties within modern societies would have to cooperate, which is rather unrealistic in such an overpopulated, impersonal and distrustful world as ours. Through this thesis Persson and Savulescu again show that liberal democracies display one of the main sources of problems modern societies have to face. Further in the book, the authors give solutions, which could help overcome the problems liberal democracies cause. Here politics would play a big role, as politicians would be the ones proposing and ordering restraints and sacrifices. But solutions such as for example cutting down birth-rates are very unlikely to be accepted especially by western democracies. It seems rather difficult to find a policy which would be accepted by the affluent countries, who are used to a very high standard of living, and simultaneously improving the condition of our planet. Consequently, as democracies seem to be unable to solve the problems displayed, the next step the authors introduce is a development to dictatorships. These would be able to accomplish political actions and vast changes within short periods. Persson and Savulescu say that the liberty existing in modern democracies nowadays will sooner or later endanger the persistence of humanity.

In the very last chapter of the book, the authors present a solution for the 'ill-equipped' humanity. They therefore present a strategy called moral enhancement. Up until now moral problems liberal democracies had to face were solved through international organizations like the United Nations. It is Persson and Savulescus point of view that in order to achieve changes, the voters of liberal democracies need to undergo a moral enhancement. For this, the authors argue that the combination of moral bioenhancement, in which patients incorporate drugs in order to make them act morally, and traditional moral education presents a possible solution. Concerning the treatment with drugs, Persson and Savulescu concentrate on two active ingredients, oxytocin and serotonin, which manipulate biology in order to evoke moral effects on the human mind. This moral bioenhancement is seen as an enhancement to 'common-sense morality'. It is the authors' point of view that in order to use the largely progressed technology in a moral way, a moral enhancement seriously needs to be taken into consideration. The book ends with Persson and Savulescus proposition of moral bioenhancement as a possible solution for future challenges and problems.

Concerning this moral bioenhancement, there are a few arguments underlining a negative outcome towards the authors' proposition. I would like to present these in the following.

The authors talk about functioning systems and say "in order to improve its function, we have to discover a condition which fits in so well with all these conditions that the function is enhanced" (p. 13-14). Again this underlines the point that it is easier to harm a well-functioning system rather than do good, heal or repair it. In this sense, as we don't know how our system would react to a moral bioenhancement, it is just as likely we harm the system as it would be to repair, or even enhance it. Finding the puzzle piece improving our system is a rather difficult task and will take great effort. As it is harder to improve the condition of the harmed system, it would tremendously very difficult to find the fitting component in order to reset or even improve the status of a harmed system. It seems to be a matter of impossibility.

Another point I would like to discuss is the determination coming alongside with the moral bioenhancement. Concerning this point, the authors say that the moral bioenhancement made possible through a drug people would have to incorporate, extends these peoples' freedom. Their thesis is that "when we influence the motivational states of people, this could be liberating rather than constraining" (p. 114). Indeed, if people choose to undergo the moral bioenhancement and incorporate the drug through their own and free will, this leads to the assumption they are not determined within their freedom. Still the drug constitutes a chemical substance influencing our system from outside of our body and system. This, in my opinion is indeed a kind of determination, especially when this has an influence on evaluating, choosing and performing our actions. To go further at this point, through this moral bioenhancement, liberty, not only within us, but also within the liberal democracies would be diluted. This is because the citizens of these liberal democracies are no longer liberal, if their actions are under chemical influence from the outside. They are determined within their choices and activities, making a completely honest living amongst each other near to impossible. Through this, again liberal democracies would fail to hold their main concept, namely granting liberty and freedom for all citizens.

Overall one needs to keep in mind that, as the authors say "in order to improve its function, we have to discover a condition which fits in so well with all these condition that the function is enhanced"(p. 13–14). Thus, taking away something of a system is always way worse than enhancing it through something. In this sense, Persson and Savulescus idea of enhancing morality and through this our society, is a good approach to solve problems future generations will have to face. It at this point is questionable whether a

bioenhancement is the right way, but the general proposition of an enhancement in my opinion leads towards the right direction.

REFERENCES

Howard-Snyder, F. (2011). Doing vs. Allowing Harm. In Zalta, E.N. (ed.), *The Stanford Encyclopedia of Philosophy.* http://plato.stanford.edu/archives/win2011/entries/doing-allowing [Accessed on 05/03/14]

Persson I., & Savulescu J. (2012). *Unfit for the Future.* New York: Oxford University Press.

Book Review
Körper 2.0

by Karin Harrasser, Bielefeld, Transcript Verlag, 2013

Matthias Moosburger [†]
moosburger@lrz.uni-muenchen.de

Top-class sport prosthesis, individually adapted running shoes and Google glasses are only a few examples of several cultural symbolisations that seem to suggest the increasing presence of the bodies nowadays which do not simply conform to organic originality but urge to be conceived as "bodies 2.0". In her recent essay *Körper 2.0* the professor for cultural studies at the University of arts in Linz, Karin Harrasser, gets to the bottom of enhanced, optimised and marketed bodies that witness the symbiotic coexisting of technology and/on the human body.

Karin Harrasser only at first glance seems to follow the German philosopher Peter Sloterdijk, who recognises these body modifications and optimisations as "anthropotechnics" differentiated in strategies of immunisation and enhancement. Whereas the former serve the viability of deficient beings, and the latter meet a "vertical tension" and longing for self-transformation and perfection, both interpretations of anthropotechnics for Karin Harrasser are obliged to a "too strongly modernistic" narrative of loss (p. 11). She in contrast tries to situate current entanglements of logics of improvement, technologies and human bodies as consequences of historical, epistemological and political highly pre-conditionally dependent assemblages. By portraying two athletes, the world record holder and currently in media present Oscar Pistorius and the athlete, actress and fashion model Aimee Mullins, who both have been amputated their legs lower the knees in childhood due to a genetic malformation, Karin Harrasser illustrates her thesis: the development in anthropotechnics corporated here, exemplified by those, who the British TV broadcaster "Channel 4" promoted as "the Superhumans" at the Paralympic Games in London 2012, not merely reveals an "expression of a

[†] Ludwigs-Maximilians-Universität München, Germany.

inner culture of self-improvement". This phenomena refers to an entanglement with a "neocapitalist logic of self-optimisation" that produces "privileged, diverse-corporal" role-models following the motto:

> You just have to imagine or wish your body in one way or another – namely stronger, faster, more beautiful, more mysterious –, it yet transforms into it.

Following this logic corporal adversities aren't disclose limitations any more, but challenges that evoke militancy.

Harrasser tries to situate her essay in between this dubious optimism on technological body enhancement as well as an exaggerated hostility towards technology on the other hand. She finds argumentative support in Donna Harraway and proclaims that "the exploration of technology's and fiction's transformative potential doesn't inevitably leads to omnipotent fantasies like trans-humanism, crypto technics and space conquering" (p. 13). She tries to diffuse the logic of alienation and optimisation, and suggests another story to be told on corporality, acting as a product and setting of multiple negotiating practices instead of mourning the loss of an isolated, omnipotent [male] subjectivity, that has never existed anyway.

Here Harrasser's theory of partially sovereign bodies draws on. She understands body technologies as "artefacts that animate our imagination and have already transformed bodies in another"(Harrasser, 2013 p. 103). These biotechnologically hybridised bodies have to be recognised as an assemblage, "that is what it is, but could be any different as well" (ibd., p. 73). Therefore she develops a perception of technology that does not simply represent "neither capitalism's machines of dominance or added value nor the material basis of an endlessly reconfiguring promise of the future. They are rather embodiments of past and current relations and world generating milieus" (ibid., p. 103 f.).

Since she analyses "technological body processing out of an horizon of re-evaluation of what matters as living" (ibid., p. 87) Karin Harrasser consequentially calls for a disciplinary opening to encounter the entanglements of self-techniques, productivity and a graduation of life's value in prospect of health-care policies (Diekämper, 2013 p. 3). And with Donna Harraway she promotes a "parahumanism" (ibid., p. 103 f.) that spots a potential for emancipation in the interactivity of organic and technological agents, without misrecognition of implicit menaces.

Partially sovereign bodies do not inevitably *own* agency that implies ethics of reduction and redemption. Karin Harrasser emphasises corporal agency as being always "partially sovereign agency, always situated, never scalable or generalizable"(Harrasser, 2013 p. 119). Nobody decides confidently to which bundle oneself is associated. Since every acting is accompanied and generated by both, active and passive procedures, Harrasser and Harraway argue against relativism in knowledge and acting but rather in favour of a "strong form of responsibility, that further extends on what is not seen from a momentary point of view" (ibid., p. 125).

By finally refusing the initially given question on bodies 2.0, Harrasser disbands the idea of an open step ladder towards perfection and the forecast to any next step of improvement. But she confirms body technology's collaboration on "partial sovereignty, an on-going amalgamation and complication of agency"(ibid., p. 73). And this denial of narratives of sovereign agency gives potential to an opposition.

REFERENCES

Diekämper, J. (2013). Beam me up! Oder: Also sprach der Robotor?. pop-zeitschrift.de.   http://www.pop-zeitschrift.de/2013/12/02/beam-me-up-oder-also-sprach-der-roboter-rezension-zu-rezension-zu-karin-harrasser-korper-2-0-uber-die-technische-erweiterbarkeit-des-menschenvon-julia-diekamper2-12-2013. [Accessed on 07/03/14]

Haraway, D. (1991). A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century. In Haraway, D. *Simians, Cyborgs and Women. The Reinvention of Nature*. New York: Routledge.

Harrasser, K. (2013). *Körper 2.0. Über die technische Erweiterbarkeit des Menschen*. Bielefeld: transcript Verlag.

Book Review
Wittgenstein in Exile

by James C. Klagge, Cambridge: MA, MIT Press, 2011

Marco Nuzzaco[†]
marco.nuzzaco@uniba.it

The work of James Klagge is an interesting operation of exploring Wittgenstein's philosophy. It assumes exile as the favourite point of view of looking at Wittgenstein's life and work as two sides of the same coin. The dimension of exile here is encompassed in its all manifold facets: existential, geographical, historical and intellectual. Thus Klagge draws a picture of Wittgenstein intertwining biographical moments with the nature of his philosophical work. He seems to suggest that we may get an authentic understanding of his philosophy by engaging with Wittgenstein as a distant "other", namely by seeing Wittgenstein as a spirit that felt himself as stranger within its time.

The first chapter, *No one Understand me*, describes Wittgenstein's intellectual exile. Here Klagge accurately portrays Wittgenstein's worry of *not be understood*: it was indeed a repeated obsession of Wittgenstein that his thoughts would have failed to hit upon the people who were coming across his thought. His two main works, the *Tractatus* and the *Philosophical Investigations*, do reflect this concern. A similar fear seems to have followed Wittgenstein even in his later philosophical activity. What interestingly emerges by Klagge's descriptions is Wittgenstein's perpetual idiosyncrasy of showing his work in front of a large public: the restricted number of students at his lectures and the few people truly acquainted of his work speak in favour of Wittgenstein's refusal of talking in front of a broad audience. Moreover, looking back at the genesis of the *Philosophical Investigations*, Klagge presents Wittgenstein's later philosophy as a kind of "esoteric" work (Klagge, 2013a). He suggests that even "the two voices" of the *Investigations* would hint at "a dialog" of Wittgenstein with himself that, according to Klagge, would further indicate his difficulty of sharing his work publicly. Even though

[†] Università degli Studi di Bari, Italy, and University of East Anglia at Norwich, UK.

it may be questionable thinking of Wittgenstein's later work as *esoteric*, Klagge's description sheds light on important biographical aspects often overlooked in regard of the background of his later philosophy.

Chapters 2, *Can we Understand Wittgenstein?* and 3, *What is Understanding?* focus on some of the methods of the examination of language exposed by Wittgenstein within the *Investigations*. Drawing on the influence that Spengler had had on Wittgenstein, Klagge describes the historical background of his time, in which Wittgenstein thought about himself as an exile; namely, as living in an era of a declining civilization where he thought his ideas could not flourish. Wittgenstein's concepts as *family resemblances*, *language games*, *seeing aspects* as *seeing as* are explained in light of his criticism of Platonic essentialism. Klagge takes these concepts as similar to the spirit of Goethe's morphological investigation of nature – as presented in Goethe's *Metamorphosis of Plant*: the spirit of the *Faustian era*, in which Spengler places Goethe's temperament, contrasts with the spirit of the ancient civilization of which Socrates's attitude, at least in early Plato's dialogues, would represent a sort of pre-modern prejudice of the scientific spirit of precision. Wittgenstein's critique of both the method of science and Platonic essentialism would be representative of this transitional passage from an era of "culture" to "civilization", in Spenlger's terms. Thus, Klagge places the reception of Wittgenstein's ideas in such a difference of temperament. He writes: "The most important point is that the diagnosis and the treatment that Wittgenstein finds are attitudinal and not cognitive. They have to do with our needs and the direction of our attention. It is possible to say that the difference between Socrates and Wittgenstein is a spiritual difference, and not simply an intellectual one" (p. X). Therefore Wittgenstein's philosophy is seen as contrasting not only intellectual tendencies, but as having a temperament alien to "the darkness of this time" – as Wittgenstein himself wrote in the preface of the *Philosophical Investigations*. Further on Klagge writes: "We can criticize Wittgenstein's presentation of his ideas, as himself does, but we have to recognize what it is in *us* that leads us to be unreceptive to those ideas" (p. X).

Wittgenstein's critique of language is further analysed in chapter 7, *Philosophy and Science*, in connection with science. Here Klagge remarkably distinguishes between science and "scientism". This distinction is often overlooked in regard of Wittgenstein's relation to science: namely, if Wittgenstein's attitude towards science was part of its alienation from the modern world, that does not amount to say that he fully rejected science as a

form of understanding as such. His knowledge of physical theories of Boltzmann, Hertz, Maxwell and Einstein proves the opposite. He was rather hostile to the idea of progress of the modern spirit of "civilization" that excludes other forms of understanding. In the draft for a preface to a book in 1930 Wittgenstein writes: "Our civilization is characterised by the word progress [...] Typically it constructs [...] And even clarity is sought only as a means to this end, not as and in itself. For me on the contrary clarity, perspicuity are valuable in themselves. I am not interested in constructing buildings, so much as having a perspicuous view of the foundations of possible buildings. So I am aiming at something different than are scientists & my thoughts move differently from theirs" (Wittgenstein, 1984, p. 7). Wittgenstein was contrasting then the tendency in philosophy of adopting methods of science by so reducing philosophy as a form of scientism. Thus, assuming Wittgenstein's idea of "clarity as an end in itself" Klagge takes philosophy responsible for how science changes meanings of words by introducing new concepts. Klagge attributes to Wittgenstein what he calls an *insulation thesis* for which "science is not relevant to the resolution of philosophical problems" but philosophy either is powerless for what regards the resolution of scientific problems. All that philosophy can is looking at language, calling it critically into question to understand how scientific discoveries affect the use of our concepts, to the extent that either they may extend their meanings, by so completely changing concepts itself; or if it is the case that they do not affect the meanings of our old concepts at all. Although Klagge takes the insulation thesis as widespread throughout Wittgenstein's philosophical career, he focuses more on his later work: here *criteria* have a central role to understand how meanings of concepts can change. By following Wittgenstein Klagge argues that "any change in the criteria – that fix the meaning of a concept – constitutes a change in the concept [...] Changing the criteria changes the subject" (p. X). That is the case of mental and psychological concepts, which Klagge deals with from chapter 7 to chapter 9, *Science and Mind*. Klagge applies Wittgenstein's examination of language to changes that neurosciences may operate on the meaning of mental concepts like pain sensations. Here Klagge isolates two different approaches to this problem. The first is *eliminative materialism theories*: if neurosciences can radically change the meanings of mental concepts, it will substitute them with concepts that will explain mental and psychological phenomena in chemical or physical terms. That is the case when our ordinary mental concepts can be

replaced by *technical* ones. As Klagge writes: "If our ordinary notions of mental states are eliminated, as predicted by eliminative materialists, then the problem will disappear, because the concepts that gave rise to it will have disappeared (and any parallel problem regarding the new concepts can be resolved by direct inspection of people brain's)" (p. X). That is a possibility that according to Klagge Wittgenstein's himself did not exclude: "Perhaps the concepts of such a language would be more suitable for understanding psychology than the concepts of our language" (Wittgenstein, 1986, §577). Nevertheless, I think here Klagge is misquoting Wittgenstein, because in the context of §577 Wittgenstein seems to referring to concepts of psychology more in a broad sense, rather than to scientific psychology as such or – as Klagge does – to neuroscience in particular. However, in another place Wittgenstein criticizes the eliminative materialist hypothesis in a way similar to Klagge's way of rejecting it. In *Zettel* §611 Wittgenstein writes: "The prejudice in favour of psychophysical parallelism is a fruit of primitive interpretations of our concepts. For if one allows a causality between psychological phenomena which is not mediated physiologically, one thinks one is professing belief in a gaseous mental entity". Therefore, on the other hand Klagge presents the possibility of there being *identity theories* according to which meanings of ordinary language or folk psychology's concepts can remain the same in spite of neuroscientific progress. That is expressed by Wittgenstein's saying that "we can't alter the mark of concept without altering the concepts itself" (Wittgenstein, 2005, p. 182) *but* "we don't [always] use language according to this strict rules either" (Wittgenstein, 1986, p.25). That is the case of words like "desire" in regard of the work of psychoanalysis. As Klagge suggests: "Should we, under pressure of Freud's work, for example, allow there to be *unconscious* desires? Should *those* – whatever we call them – count as desires too? But by now, if not in Wittgenstein's time, it is widely accepted in society that this makes sense. On what ground could one insist the unconscious desires is not a desire? To hold that the concept of desire *could not apply* to what Freud was talking about seems excessively *a priori*" (p. X). That is a form of what Klagge calls *compatibilism*, that is, a kind of *bilingualism* in Klagge's words, for which both concepts of neuroscience and folk psychology can coexist for different practical purposes.

However, Klagge takes Wittgenstein not as to fully reject the position of *materialist eliminativism* for the sake of *compatibilism* as such; but in face of Wittgenstein's criticism he is rather to ask "what kind of life we want to live in

face of progresses of science? Since science changes our language, our conceptual life, and therefore the ways we relate to realty, how does it affect the way we decide to live?" That is the case when Klagge writes: "The only philosophical resolutions of puzzles is by Wittgenstein's method, to which science can lend no assistance in any form. Wittgenstein articulates this point carefully in a manuscript (February 9, 1948; MS 137, 15a):

> One must just take the concept of seeing as one finds it; not want to refine it. – But really, why not. – Because it is not our problem to change it, to introduce (as science does) one adapted for some purpose or other, but to understand it; on order not produce a false conception of it.

And further on he adds: "Whether one employs current measures or advanced neuroscientific ones, however, it is not obvious that humankind will be better off by pursuing neuroscience. If compatibilism is true, do people want science to continue a policy that, by eliminativists' own admission threatens rational agency? To put the issue in Wittgenstein's terms: which form of life do we want to live in? Reflecting on this very question during his Lent term, 1939, lectures, Wittgenstein is reported as saying: "Although is conceivable that if we had a mechanism that would show all this, we would change our terminology – and say, "He's as much compelled as if a policemen shoved him". We'd give up this distinction then; and if we did, it would be very sorry". Therefore, I think the way Klagge takes Wittgenstein's methods of clarifying language is epistemological as much as ethical, to the extent that it also suggests a moral reflection. The case of medicine is a clear example. The purposes of medicine of "predict the course of suffering, prevent the suffering from spreading to others, and cure the suffering" stand beside the necessity of comforting the sufferers, of "acknowledging the humanity in a man" (Wittgenstein, 1984, p. 1), of having "an attitude towards a soul" (Wittgenstein, 1986 p. 178) as Wittgenstein would say. If our concepts might be significantly altered – as the hypothesis of eliminative materialism argues – or if they are compatible with those we already have, that cannot change the meaning of fundamental attitudes like comforting. That is something that Wittgenstein reminded to Drury when he was studying psychiatry. Remembering his experiences at the Guy's Hospital and in Newcastle as an orderly during the World War II, he said in 1948 to Drury: "Always take a chair and sit down by the patient's bedside: do not stand at the end of the bed in a dictatorial attitude. Let your patients feel that they have time to talk with

you". So here Klagge fully captures this ethical aspect of Wittgenstein's thought: "To the extent that traditional concepts of health care embody a concern for comfort and suffering, their whole purpose will not be subsumed by the concepts of scientific medicine. If they are nevertheless replaced by the concepts of scientific medicine, something valuable may be lost"(p. X).

At any rate, Klagge deals with a more ethical aspect of Wittgenstein's thought in chapter 10, *Das erlösende Wort*. The title represents Wittgenstein's attitude on which Klagge focuses: Klagge provides a collection of recurrences – from Wittgenstein's secret war diaries to his later manuscripts – in which Wittgenstein was speaking of looking for the *erlösende Wort*: he points out that this expression has been mistranslated by E. Anscombe with "key word" rather than as "the redeemer word" as it correctly should have been. If in his early work, especially in the *Tractauts*, the "redeemer word" was a claim of silence, it nonetheless has an important role in his later work. Regarding clarity as a philosophical goal *per se* Wittgenstein warned us against the hazard of not acknowledging when in philosophy clarity is achieved. Thus, the *erlösende Wort* requires, according to Klagge, Wittgenstein's temperament of realising when to stop, namely of "suppressing the urge to explain, justify" on pain of distorting clarity – as Wittgenstein himself often repeated: "It's important in philosophy to know when to stop – when not to ask a question"; "The difficulty here is: to stop" (Wittgenstein, 1967, §314); "The real discovery is the one that makes me capable of stopping doing philosophy when I want to"; (Wittgenstein, 1986, §133). "If I have exhausted the justifications I have reached bedrock, and my spade is turned. Then I am inclined to say: "This is simply what I do." (Wittgenstein, 1986, §217). Klagge explores this attitude of Wittgenstein in three different trajectories: by addressing it to Wittgenstein's approach to *Divine Command Theories*, i.e., his ideas on moral imperative and religious beliefs; the different attitudes of Alyosha and Ivan of *The Karamazov Brothers* and his interests towards the Mormons. Although Klagge captures important aspects of Wittgenstein's attitude towards religious beliefs, in some passages it gives the impression of attributing such an attitude of Wittgenstein as something inaccessible to others devoid of his temperament. That is the case when Klagge writes: "So too with the *erlösende* word – what works may well depend on the person the person's state of mind" (p. X) or "The *erlösende* word does not work in the face of all temptations – in all traditions – and can only be effectively spoke under the right circumstances. It may not easily be

understood by us" (p. X). But if this is the case, Klagge's words here sound misleading. We won't get the importance of what Wittgenstein aimed to convey, namely the fact that acknowledging the *erlösende* word is an attitude which as to work as reminding us how to improve what Wittgenstein called "an attitude towards a soul", no matter in what circumstances or state of mind we might be, neither who is speaking or to whom we are speaking to. Knowing when to stop is so an attitude that has to be cultivated rather than taken for granted or as spoken on behalf of a limited number of people who shared (or might me be able to share) Wittgenstein's temperament. If that is a matter of temperament, Wittgenstein's teaching of when to stop is a task that no one can carry out before us; rather it urges us to acknowledge *the humanity in a man* behind his religious beliefs or his moral imperatives. The very fact that Wittgenstein spoke of the *erlösende* word as the need of knowing when to stop that is too something we can get from him in regard of religious beliefs and others' moral imperatives. That is even more true if in ethics we are to adopt a realist view that is of pluralism rather than simply of relativism. If this is what we can get from Wittgenstein, it does not mean that we have to share his temperament and his exile full blown.

However, Klagge himself seems to be aware of this in other places. In the last chapter, *Wittgenstein in the Twenty-First Century*, he describes how Wittgenstein's teachings have been undertaken by other philosophers until today in different and autonomous directions. He refers to the cognitive scientists Eleanor Rosch and Carolyn Mervis, "who undertook empirical investigations of categorizations using the structuring of Wittgenstein's family resemblances", to Putnam's mental experiments about construal of mental states in regard of meaning and to constitutivism's theories like Quine and J. Kim's concepts of supervenience. These are meant to be a sign, in Klagge's view, of how Wittgenstein's ideas can be of a great help within present debates, despite the fact that "something was obstructive in his temperament – [despite] his own difficulties of engaging with people" (p. X). Klagge traces such a difficulty back to his condition of exile explored throughout his book. How much his exile influenced his temperament is difficult to say, no more than judging in light of other relevant factors such as his personal considerations of "how best to live" and "what was most important for him". I do think these considerations, often repeated by Wittgenstein himself, can be part of his temperament as much as his exile, without which his ideas would not have had the impact they still have on philosophy.

At any rate, if we accept Wittgenstein's exile as a key to approach his philosophy – as presented by Klagge's book – we should also resist to the temptation of not being much caught in Wittgenstein's temperament, if that might obstruct any possible cultivation of our own temperament. Once again, that is the best reminder that Wittgenstein has to give us. I read in this way one of Klagge's final remarks: "No doubt exiles have difficulty engaging with those around them. But that does not mean those around them have to suffer from the same difficulties. Perhaps Wittgenstein helps us to proper appreciate the differences there are between people and not to underestimate the difficulties. But in this century we hardly need reminding of that".

REFERENCES

Klagge, J.C. (2010). *Eliminativism and the Mental. Epistemologiia & Filosofiia Nauki, 24*(2), 17–27.

Klagge, J.C. (2010) (ed.). *Wittgenstein. Biography and Philosophy.* Cambridge: Cambridge University Press.

Klagge, J.C. (2013). *Wittgenstein in Exile.* In Phillips, D.Z., Rhees, R., & Von der Ruhr, M. (eds.), *Religion and Wittgenstein´s Legacy.* Aldershot: Ashgate.

Klagge, J.C. (2013a). *Wittgenstein and His Audience: Esotericist or Evangelist?* in *The Textual Genesis of Wittgenstein's Investigations*, ed. N. Venturinha, Routledge, London.

Wittgenstein, L. (1961). *Tractatus Logicus Philosophicus.* London: Routledge.

Wittgenstein, L. (1967). *Zettel.* Oxford: Basil Blackwell.

Wittgenstein, L. (1984). *Culture and Value.* Chicago: Chicago University Press.

Wittgenstein, L. (1986). *Philosophical Investigations.* Oxford: Basil Blackwell.

Wittgenstein, L. (1993). *Philosophical Occasions.* Edited by Klagge, J.C., & Norman, A. Indianapolis: Hackett.