

THE PERILS OF COGNITIVE ENHANCEMENT AND THE URGENT IMPERATIVE TO ENHANCE THE MORAL CHARACTER OF HUMANITY

Ingmar Persson and Julian Savulescu

As history shows, some human beings are capable of acting very immorally¹. Technological advance and consequent exponential growth in cognitive power means that even rare evil individuals can act with catastrophic effect. The advance of science makes biological, nuclear and other weapons of mass destruction easier and easier to fabricate and, thus, increases the probability that they will come into the hands of small terrorist groups and deranged individuals. Cognitive enhancement by means of drugs, implants and biological (including genetic) interventions could thus accelerate the advance of science, or its application, and so increase the risk of the development or misuse of weapons of mass destruction. We argue that this is a reason which speaks against the desirability of cognitive enhancement, and the consequent speedier growth of knowledge, if it is not accompanied by an extensive moral enhancement of humankind. We review the possibilities for moral enhancement by biomedical and genetic means and conclude that, though it should be possible in principle, it is in practice probably distant. There is thus a reason not to support cognitive enhancement in the foreseeable future. However, we grant that there are also reasons in its favour, but we do not attempt to settle the balance between these reasons for and against. Rather, we conclude that if research into cognitive enhancement continues, as it is likely to, it must be accompanied by research into moral enhancement.

1. INTRODUCTION

Enhancement of our capacities or faculties is *instrumentally* good for us when and because it enables us to realize our ends more effectively and reliably. This is true of both enhancement of physical capacities, like the strength of our bodies or their resistance to

disease, and enhancement of mental capacities, including cognitive and affective capacities. But we shall here focus on enhancement of cognitive faculties because it offers the greatest possibility of making us better in attaining our ends. For instance, by enabling us to make various technical inventions, cognitive enhancement allows us to get around our physical limitations.

What is instrumentally good relative to a certain end may be bad, however, rather than good because the end itself is bad rather good. Among your ends, we could distinguish roughly between ends which are good for you, or *prudentially* good, and good for everyone on the whole, or *morally* good. Cognitive enhancement is prudentially good for you when it makes you more proficient in pursuing your individual prudential ends and helps you ensure that your individual ends are internally consistent and, so, that your fulfilment of ends could be as great as possible. Cognitive enhancement is prudentially good for agents when it improves their own well-being or quality of life.

What is prudentially good for you may, however, be prudentially bad for others, for your success in respect of fulfilling your prudential goals may make it more difficult for others to fulfil their prudential goals. Fortunately, this is not necessarily so. It is not the case if your prudential ends, as well as the prudential ends of others, are sanctioned by morality. For it is a familiar idea that it is definitive of morality to harmonize the prudential ends of people, so that they be co-satisfiable. If, however, there is a clash of prudential ends, because to a greater or lesser degree people have immoral prudential ends, your cognitive enhancement is likely to be bad news for others, just as their cognitive enhancement is likely to be bad news for you, since it will undercut the boost of your prudential success or well-being that your cognitive enhancement by itself promises.

Thus, if enhancement of our capacities to achieve our ends is to be for the better for all on the whole, our ends must be in accordance with the demands of morality. Generally speaking, the greater the ratio of moral people to immoral people, the better overall is

enhancement of our capacities to achieve our ends. But we shall argue that, even if almost all of us are moral to a high degree, there is a good reason to think that a cognitive enhancement, and a consequent faster increase of knowledge, which extends to all of us may be worse for us on the whole than no cognitive enhancement at all, if there is a minority which is morally corrupt. This may appear surprising at first glance, for it may be thought that, if almost all of us are moral, or aim for what is best for all of us on the whole, cognitive enhancement of all of us which makes us more proficient in fulfilling our aims must be for the best of all of us on the whole. But we shall argue that this is not true in all circumstances. And even if the expected utility of cognitive enhancement outweighs its expected disutility, there may be important reasons not to pursue or employ it, reasons to do ultimately with the very survival of humanity itself.

2. COGNITIVE ENHANCEMENT: TRADITIONAL AND BIOMEDICAL MEANS

So far in history the cognitive enhancement of human beings has taken the shape of education, or knowledge which has been transmitted from earlier generations to later ones, first orally, then by writing as well. This accumulation of knowledge is what makes, for instance, the best mathematicians of today capable of figuring out solutions to mathematical problems that Euclid could not have dreamed of. Thus, in relation to Euclid, contemporary mathematicians are mathematically enhanced by the traditional method of education. There is no reason to believe that their hereditary mathematical gifts are greater than his. It appears that the human species has been essentially the same in biological and genetic respects for at least 40 000 years, that is, while the main part of the unparallel human cultural development has taken place² What has made this explosive development possible is not any radical biological or genetic change, but simply that human beings were already 40 000 years ago smart and social enough to begin to develop oral languages which enabled them to share knowledge with each other. Later, with the invention of writing, they could more effectively

transmit knowledge from earlier to later generations. This growth of knowledge provided them with means of improving their standard of living, which in turn led the human population to grow. As a consequence, there was ever more human mind power that could be pooled together through the network of language, both oral and written. In briefest outline, this is what has made possible the so far as we know unprecedented cultural, artistic and scientific development which constitutes human history.

It bears repeating that this development has occurred without any significant biological or genetic change in respect human cognitive faculties, for we are only now beginning to acquire the means of effecting such changes. Let us review these means.

SOME CURRENT POSSIBILITIES OF BIOLOGICAL COGNITIVE ENHANCEMENT³

GENETIC MEMORY ENHANCEMENT

Genetic memory enhancement has been demonstrated in rats and mice. For example, Tang et al⁴. modified mice to overexpress or overproduce a receptor in the brain called NR2B. The NR2B “Doogie” mice demonstrated improved memory performance, both in terms of acquisition and retention⁵. The modification also made them more sensitive to certain forms of pain⁶. In other experiments, increased amount of brain growth factors⁷ and proteins (such as the signal transduction protein, adenylyl cyclase)⁸ have also produced memory improvements.

MEMORY ENHANCING DRUGS

There are several groups of memory enhancing drugs affecting the learning and encoding process. They range from stimulants⁹, nutrients¹⁰ and hormones¹¹ over cholinergic agonists¹² and the piracetam family¹³ to ampakines¹⁴ and consolidation enhancers¹⁵.

A notable form of chemical enhancement is pre- and perinatal enhancement. By giving choline supplementation, which is abundant in eggs, to pregnant rats the performance of their pups was enhanced, apparently by changes in neural development¹⁶. Given the ready availability of choline such prenatal enhancement may already (inadvertently) take place.

Enriched rearing environments have been found to increase brain development and improve cognition¹⁷. There is evidence that patients with Alzheimer's Disease that another class of drugs (cholinesterase inhibitors) slow cognitive decline¹⁸ by producing a more robust and plastic neural structure.

WORKING MEMORY

Working memory can be modulated by a variety of drugs. Drugs that stimulate the dopamine system have demonstrated effect as well as cholinergic drugs¹⁹. Modafinil is a new class of drug designed for sleep disorders but has been shown to increase alertness and attention and to enhance working memory in healthy test subjects²⁰. It was also found to enhance digit span, visual pattern recognition memory, spatial planning and reaction time/latency on different working memory tasks²¹.

In addition, there have been promising experiments and analyses involving procedural memory, cortical reorganization, creativity, executive function and human computer interaction which could serve as a basis for cognitive enhancement²².

SELF-CONTROL

According to the “ego depletion” theory, exerting self-control consumes a limited resource, reducing the amount available for subsequent self-control efforts. Various experiments have documented that previous exercise of self-control (including coping with stress, suppressing aversive thoughts and resisting temptation) increases the likelihood that subsequent weakness of will, despite not affecting performance of behaviors that do not

require self-control²³. The effect might also extend to other forms of volition²⁴. While this resource has mainly been seen as a metaphor, recent studies suggest that at least part of it may be energy in the form of blood glucose, and that consuming glucose drinks after depleting tasks improves self-control during a subsequent task²⁵. If these results hold it appears likely that pharmacological or physiological interventions (to improve glucose regulation) may be used to improve self-control abilities.

MENTAL ENERGY AND WAKEFULNESS

Great achievers are in general not just equipped with great cognitive abilities but also an abundance of mental energy²⁶. Aspects of this energy are the ability to focus on a problem for a long time, a strong motivation to succeed, the ability to stay awake and work hard. It might also be a temperament²⁷, potentially linked to hypomanic or manic states. While mental energy borders on mood enhancement rather than cognition enhancement, it is clear that interventions improving it will have enabling effects on other cognitive capacities. Stimulants, such as modafenil, may be able to produce more mental energy. The US military routinely uses both modafenil and Ritalin for its pilots.

New sleep- and awakesness-controlling medications such as modafenil appear to promote heightened function with small risks of direct side effects and dependency²⁸. The effects of modafenil and amphetamine on sleep-deprived aviators were found to be similar, but with modafinil showing fewer side effects²⁹. Naps are more effective in maintaining performance than modafinil and amphetamine during long (48h) sleep deprivation than short (24h), and naps followed by a modafinil dose may be more efficient than either individually³⁰. The drug also improved attention and working memory in sleep-deprived physicians³¹. Modafenil also allows sleep when appropriate and has frontally acting cognition enhancing effects (see above).

3. THE DANGER OF COGNITIVE ENHANCEMENT AT THE PRESENT TIME

There is thus a considerable potential for a significant improvement of various aspects of human cognition by biomedical means. We have considered internal biological enhancements, but we must not forget external aids that also serve to enhance our cognitive processing, such as access to supercomputers. This avenue may give us access to cognitive powers which otherwise would be beyond our reach. In the near future, access to vastly enhanced computing power and connection of minds and information through the internet seem the most realistic means of substantial cognitive enhancement.

It is likely that substantial cognitive enhancement will speed up the acquisition of knowledge. If, say, the mathematical gift of Euclid and subsequent mathematicians had been biomedically or genetically enhanced, the human species would presumably sooner have reached the state of contemporary mathematics and, consequently, hypothetical present day mathematicians would have reached a state of art that actual present day mathematicians cannot divine. But we should not assume that a genetically or biomedically enhanced contemporary mathematician would have grasped truths that unenhanced humans could never grasp, whatever the period of time at their disposal. For all we know, it may just be that *more* generations of unenhanced mathematicians are required to arrive at these truths. Thus, the most likely effect of cognitive enhancement by genetic or biomedical means may be to speed up a growth of knowledge that would otherwise have taken humanity a longer time to achieve.

The question we want to ask is whether such a speeding up of the growth of scientific knowledge is desirable at the present point of time, or in the near future. It is well known that computing power continues to double every 2 years – Moore’s law. So at the very least, computing power, and the cognitive enhancement it affords, and other cognitive enhancements will increase at an exponential rate. We want to call attention to one reason for thinking that this is not desirable. This reason is that this expansion of scientific knowledge

and cognitive ability will put in an increasing number of people's hands "weapons of mass destruction" or the ability to deploy them. In so far as this is so, this growth of knowledge will be instrumentally bad for us on the whole, by unacceptably increasing the risk that we shall die soon. It will be bad for us that scientific knowledge continues to grow by traditional means, and even worse if this growth is further accelerated by biomedical or genetic enhancement of our cognitive capacities. For if an increasing percentage of us acquires the power to destroy a large number of us, it is enough if very few of us are malevolent or vicious enough to use this power for all of us to run an unacceptable increase of the risk of death and disaster. To eliminate this risk, cognitive enhancement would have to be accompanied by a *moral* enhancement which extends to *all* of us, since such moral enhancement could reduce malevolence.

A further expansion of scientific and technological knowledge — let alone an acceleration of this expansion by novel means — is problematic because we are already on the brink of acquiring — if we have not already acquired — knowledge which enables small groups, or even single individuals, to kill millions of us. Nuclear weapons are one familiar example, which have been feared since the fifties. To make a nuclear bomb out of enriched uranium may not be beyond the capacity of a well-organized terrorist group. If set off in the middle of a mega-city, such a bomb could kill millions of people and cause panic among billions, causing direct medical disaster and indirect social catastrophe. Enriched uranium, for instance from the former Soviet Union, may already be in the hands of some terrorist groups. If it is not, it could easily become so in the imminent future, when their knowledge of how to make use of it could be expected to be even greater. It is of course much harder to survey and control small terrorist groups than much larger "rogue" states. So, though the advance of science would predictably equip us with more effective methods of surveillance, even assuming that these are only used for benign purposes, this may not be enough to maintain a reassuring level of security.

Another, even scarier threat is biological weapons. These are scarier because they are even easier to construct. Diseases caused by these means could have spread widely before they are discovered because their incubation time is a week or more. This is true of small pox which kills one out of three infected. Biological weapons are also harder to control and outlaw because they are the downside of research which has the wholly commendable aim of curing diseases. Polio virus has now been artificially constructed using commercially available DNA sequences. More frighteningly, scientists have modified mousepox to make it lethal in 100% of mice. Mousepox is similar to human small pox. Voltaire estimated that small pox killed around 20% of the French population in his day. It was eradicated last century from the globe by vaccination. The study of mouse pox was published on the internet, making it indiscriminately available. Genetic engineering of small pox could create a new strain which would wipe out all or most of humanity.

We may not have yet reached the state in which a single satanic character could eradicate all life on Earth, but with cognitive enhancement by traditional means alone, we may soon be there. Martin Rees speculates about the possible invention of nanomachines which are designed to be more omnivorous than any bacterium and which could consume all known living organisms³². With biomedical and genetic enhancement of our cognitive powers we may be even closer to the invention of such monsters or other as yet unimagined demons.

The famous Cambridge physicist, Stephen Hawking, who is paralysed by motor neuron disease, conducted an unusual survey on the internet in 2006. He asked the world: what are the greatest threats to the survival of humanity and what can we do about them so that the human race can survive the next century? The poll ran for several months and people sent in all sorts of idiosyncratic suggestions. At the end of the poll, however, Hawking, one of the world's greatest intellects, gave his own answers to his questions. He argued that humanity still faced extinction, as it always has, from cosmic events such as asteroids striking earth. But climate change is a newcomer among threats to our survival. Similarly, threats from new

pathogens, such as viruses and bacteria, have increased through intercontinental travelling. Hawking concluded: "There's a sick joke that the reason we haven't been visited by aliens is that when a civilisation reaches our stage of development, it becomes unstable and destroys itself...The long-term survival of the human race will be safe only if we spread out into space, and then to other stars. This won't happen for at least 100 years so we have to be very careful. Perhaps, we must hope that genetic engineering will make us wise and less aggressive." ³³

The human species, and the rest of life on Earth, faces a series of disastrous threats. Some of these have been with us throughout our whole history. Scientific progress has helped us to protect ourselves against some of these, and will probably help us to protect ourselves against more in the future. But the irony is that the very same progress creates other equally lethal threats. It is hard to tell how this balance comes out, whether the chances of humanity to survive and prosper in the future are increasing or decreasing at present. Whatever the balance, in order to reduce the existential threat that cognitive enhancement poses, we require a moral enhancement, an enhancement of our motivation to act morally. The threats come not only from cognitive enhancement by novel biomedical and genetic means, but also from the growth of knowledge by traditional cultural means, and by external means, such as access to supercomputers. Indeed, it may be that we already are too cognitively advanced for our own good (and for the good of other species on Earth).

4. THE POSSIBILITY OF MORAL ENHANCEMENT

Since we can hardly intentionally retrace our cognitive steps, erase from our collective memory what science has learnt about the workings of the world, or even stop further scientific progress, especially progress in computing, it seems likely we will continue to radically enhance our cognitive powers at an exponential rate. Let us then ask what the prospects of moral enhancement are. It is obvious that moral enhancement by traditional,

cultural means — i.e. by the transmission of moral instruction and knowledge from earlier to subsequent generations — has not been anything like as effective and quick as cognitive enhancement by these means. One reason for this is that, contrary to what Socrates is reputed to have believed, we do not necessarily do what is right and good as soon as we gain knowledge of what this is. We have made some moral progress owing to cultural forces, although moral knowledge has turned out to be harder to come by than scientific knowledge (this is another reason why moral progress has been less impressive than scientific progress). For instance, the realization of the falsity of racism — i.e. the doctrine that the mere fact that human beings belong to different races in itself is a ground for a difference as regards their value — is of this kind. But the mere realization that racism is false is not enough to wash away all xenophobic reactions in our nature. Racial differences signify lack of kinship, and mark off strangers from neighbours. People encode the race of each individual they encounter, and do so via computational processes that appear to be both automatic and mandatory. Encoding by race is a byproduct of cognitive machinery that evolved to detect coalitional alliances³⁴. If genetic and biomedical means of enhancement could counter such natural tendencies, they could have a crucial role to play in improving our moral character, that could complement traditional social and educational means of moral enhancement. It should be stressed, though, that they are not the only means of moral enhancement, since we can evidently become morally better by training and educating ourselves. But this procedure probably requires that we are already to a significant extent morally motivated, so it is likely to be ineffective in the case of those who are wholly morally depraved or corrupt. Furthermore, these means often operate comparatively slowly if their effectiveness can be measured by how little moral progress there has been, say, in the last two thousand five hundred years — since men like Confucius, the Buddha and Socrates appeared — compared to the scientific progress in the same period. Biomedical and genetic means may be much

more effective in terms of both how thoroughly and quickly they could improve everyone in need of improvement. At very least, there is a moral imperative to explore them.

But what dispositions of ours should be enhanced to morally enhance ourselves? What is the core of our moral dispositions? And is there any reason to believe that this core is malleable by biomedical and genetic means? As we will argue, there is reason to believe that the core of our moral motivation is able to be shaped by these means, and not only by traditional cultural means, because we share this core with non-human animals from which we have evolved. So, our moral dispositions are based in our biology. They are not a cultural product to the same extent as the understanding of a language or of the laws of a society.

Needless to say, there are many accounts of what morality is, or of what it is to be moral. We cannot hope to be neutral between all these views, but the assumptions we shall make are shared by many. According to our preferred view, the core of our moral dispositions comprises, first, a disposition to *altruism*, to sympathize with other beings, to want their lives to go well rather than badly for their own sakes. Few would deny that this disposition is central to morality. It is reflected in the Golden Rule of Christianity and corresponding prescriptions of other world-wide religions, like Confucianism and Buddhism. Furthermore, philosophers like Arthur Schopenhauer thought that altruism in the shape of compassion (“Mitleid”) was *the* ground of morality. In David Hume’s terminology, “sympathy” played a similar central role in morality.

Secondly, there is a set of dispositions from which the sense of *justice* or *fairness* originates. The most basic of these dispositions are, we believe, the ones that have been called “tit-for-tat”. Evolutionary theorists have found that collectives in which this pattern of reciprocal reactions is widespread are most successful in terms of survival and reproduction³⁵. Suppose that someone does another party a favour out of altruism. Then, according to tit-for-tat, the latter should respond with *gratitude*, and a desire to return the favour with a proportionate favour. On the other hand, if someone harms another, the proper response of the offended

party is *anger*, and a desire for retaliation. This is also, to a lesser degree, the proper reaction if no, or too little, gratitude is forthcoming in response to a favour done. It is easy to see the usefulness of these responses in populations in which they are widely, but not completely, pervasive, for (adequate) gratitude encourages the giving of new favours, and (proportionate) anger discourages future aggression.

More sophisticated emotional responses which are parts of tit-for-tat are *remorse* and *feelings of guilt* if you have acted wrongly by harming someone without good reason, *shame* if you are less successful than others in returning favours, or retaliating wrongs, *pride* if you are more successful than others in these respects, *admiration* and *contempt* for others who are successful and unsuccessful in these respects, and *forgiveness* when you realize that someone is not responsible for some wrong done, or shows remorse for it. Again, it is easy to see the usefulness of these emotions. For instance, if you forgive those who are not responsible for wrongful behaviour, or who feel remorse for it, you will not waste any anger upon individuals who need not be discouraged from future aggression. But, to repeat, the usefulness of the tit-for-tat strategy depends upon it being widely, but not completely, spread in a population. If everyone else was completely selfish, it would be self-destructive to engage in the initial acts of altruism. On the other hand, if everyone was uniformly altruistic and trustworthy, reactions such as anger would be superfluous. In human populations, however, there seems to be the mixture of moral goodness and badness which makes tit-for-tat optimal.

Supposing that the dispositions of altruism and justice, whose elementary form is tit-for-tat, constitute the centrepiece of our morality, moral enhancement will consist in strengthening our altruism and making us just or fair, i.e., properly grateful, angry, forgiving, etc. More altruism is likely to initiate more tit-for-tat exchanges, though too much altruism may be an obstacle by making us turn the other cheek when tit-for-tat requires retaliation³⁶. Too little gratitude may provoke anger and aggression in benefactors rather

than further favours; too much anger in response to aggressors may spark off an escalation of violence rather than simply deterrence of future violence, and too little anger might not be a sufficient deterrent; the same is true of too little and too much forgiveness. So, altruism and these tit-for-tat emotions need to be properly attuned to be maximally useful.

Given that we now have an outline of what enhancement of moral motivation could consist in, is there any reason to believe that it can be affected by altering our biology or genes? We think there is because there is reason to believe that we share the core of moral dispositions that we have identified with some non-human animals. The biological background of altruism or unselfish behaviour has been thoroughly explored by Elliot Sober and David Sloan in their book *Unto Others*³⁷. What we mean by altruism should not be confused with mere emotional contagion — for example, fear spreading through a herd — or the distress and helping behaviour elicited by nothing but the *outward* signs of another's suffering. We mean a more sophisticated reaction which presupposes awareness that the another party, which is like oneself in possessing a mind or consciousness, is experiencing pain or suffering that is expressed by its overt behaviour. Certainly, it is difficult to tell that an animal is capable of altruism in this sense, but it seems likely that it is if it exhibits helping behaviour tailored to the individual needs of another when these are different from its own needs, such as when a chimpanzee helps a bird to fly (a chimpanzee presumably never experiences any need or desire to fly). It is believed that, alongside humans, at least apes and dolphins are capable of altruism in this sophisticated sense.

The occurrence of the tit-for-tat strategy in animal populations has also been documented. For instance, Frans de Waal has found that, among chimpanzees, “adults were likely to share food with individuals who had groomed them earlier”³⁸. This looks suspiciously like gratitude. In another of his studies, capuchin monkeys were paired with a group mate and their reactions were watched when their partner received a better reward for doing the same bartering task. The different rewards consisted in two kinds of token which could

immediately be exchanged for more tasty food, e.g., a grape, and less tasty food, e.g., a piece of cucumber, respectively. de Waal reports:

Individuals who received lower value rewards showed both passive negative reactions (e.g., refusing to exchange the token, ignoring the reward) and active negative reactions (e.g., throwing out the token or the reward). Compared to tests in which both received identical rewards, the capuchins were far less willing to complete the exchange or accept the reward if their partner received a better deal ...Capuchins refused to participate even more frequently if their partner did not have to work (exchange) to get a better reward but was handed it for "free"³⁹

De Waal concludes: "Capuchin monkeys thus seem to measure reward in relative terms, comparing their own rewards with those available and their own efforts with those of others"⁴⁰. He stresses that the reactions of the capuchin monkeys were rather "egocentric"⁴¹ (p. 49) in the sense that they reacted negatively only when they themselves were treated worse, not when their partners got the worse deal and that it therefore may not be accurate to speak of fairness without qualification. A sense of fairness in a fuller sense reasonably presupposes an ability to empathize, to put oneself imaginatively into the place of another, as does genuine altruism. But de Waal also points out that this shortcoming should not efface the continuity between these monkey reactions and a human sense of fairness or justice.

There are, however, other animal experiments whose results seem to contradict this claim by indicating that not even our closest non-human relatives, chimpanzees, have a sense of fairness. Keith Jensen and colleagues played the so-called ultimatum game with chimpanzees⁴². In this game two players, a proposer and a responder divide a reward, in the present case raisins. Ten raisins are divided between two pots in different ways, e.g., five and five, or eight and two. The chimp which is allotted the role of proposer can choose one of two trays, each with two pots with a different distribution of raisins, by pulling it halfway to its cage by means of a rope. The responder is then in a position to choose whether to pull the

tray close enough for both players to reach one pot each, or to let the tray remain beyond reach. It was found that chimp responders generally accepted 2/8 distributions without any sign of dissatisfaction even when there was an equal distribution of five raisins in each pot on the alternative tray. In contrast, under similar conditions adult human responders as a rule respond by rejecting the offer, thereby forgoing a smaller reward in order to punish the proposers for their blatant unfairness.

This result certainly indicates that chimps are blind to unfairness *in the ultimatum game*, but it would be illegitimate to infer that they do not have any conception of unfairness, that is, could not detect unfairness in *any* type of situation, e.g., the tit-for-tat situation in which we have suggested that 'elementary' manifestations of the sense of unfairness should be sought. Jensen et al seem to acknowledge this possibility when they note: 'The fact that chimpanzees in this study did not punish other individuals for making unfair offers may be in part a reflection of the fact that active food sharing is rare in this species' (p. 109). Perhaps the chimp responders saw the proposers as having a greater right to the raisins in virtue of having made the first move to get them. The situation would then be interpreted as analogous with ones which one chimp in the wild has found food and another merely gets the spoils.

Consequently, we do not think that these findings about chimps playing the ultimatum game provide any evidence for rejecting the hypothesis that the sense of fairness or justice has a biological or genetic basis. On the contrary, this hypothesis has been confirmed by another experiment in which human twins played the proposer and responder roles of the ultimatum game. Björn Wallace and associates have found that in the case of identical twins (who share the same genes) there is a striking correlation between the average division with respect to both what they propose and what they are ready to accept as responders. There is no such correlation in the case of fraternal twins⁴³. This certainly indicates that the human sense of fairness has some genetic basis.

Even if the core moral dispositions of altruism and fairness have a genetic or biological basis, it might still be thought that enhancement as regards them is not all the moral enhancement we need. For people can cause great havoc with potentially megadestructive technologies not only out of such blatantly immoral dispositions as malevolence and excessive aggressiveness. They could also do so, albeit not intentionally, through *negligence*. Now, cognitive enhancement would probably reduce the scope of negligence. To be negligent in acting consists, at least in part, in acting in spite of having failed to retrieve available relevant information about the consequences of what is being done, or having failed to process this information rationally, e.g., making the logically requisite inferences. Cognitive enhancement will reduce the incidence of such failures and will thereby increase the probability that we do what we should do⁴⁴. We are more likely to succeed in doing what we ought to do if, owing to cognitive enhancement, it is easier for us to retrieve available information and to rationally process it.

All the same, even in an enhanced state, it will take *some* effort to retrieve morally relevant information that you may fail to undertake if you lack the moral virtue of *conscientiousness*. Thus, it would seem that we should add the acquisition of this trait to what moral enhancement should comprise. Conscientiousness belongs to a category of might be called *executive* virtues. Other members of this category are strength of will or character, courage and temperance, and corresponding vices, which, alongside negligence, comprise recklessness, rashness, weakness of will, cowardice and intemperance. Unlike altruism and gratitude, these virtues are not *essentially* moral, or distinctive of a morally good person, for criminals could exhibit conscientiousness, strength of will and courage (though some people mistakenly assume that acts of terrorism must be “cowardly”).

However, we believe that executive virtues do not need separate treatment. This is because they go with strength of motivation. Thus, if you increase the altruistic motivation of people, you decrease the risk that they will negligently fail to consider the possible harmful

effects of their behaviour on their fellow-beings. Being concerned about avoiding such risks is part of what having altruistic concern for these beings consists in. Moreover, the advance of technology will in all probability bring along more effective mechanisms of surveillance, and it is easier for these to pick up people who are negligent rather than evil-doers who are intent on beating them.

So, having suggested that the core moral dispositions of altruism and a sense of fairness have a biological basis and, thus, in principle should be within the reach of biomedical and genetic treatment, the next question in line is to what extent such treatment is possible in practice. To this question the answer seems to be: only to a very small extent. A lot more scientific research is needed before we can be made more altruistic or just by suitable drugs or surgery, or genetic manipulation. Most progress may have been made in the area of anger impediment. In his paper in this volume Tom Douglas refers to some relevant findings:

- Oxytocin has been shown to promote trust
- SSRIs to increase co-operation/reduced aggression

Ritalin has been given to children with Attention Deficit Disorder reduces violent aggression
In addition, the biological basis for some personality types which predispose to immoral behaviour is beginning to be elucidated. For example,

- Antisocial personality disorder may have a biological basis
- Criminality has been linked to MAO mutation on the X chromosome, especially when coupled with social deprivation.

As these conditions are better understood, it is possible that interventions will be developed which improve behaviour.⁴⁵

Our knowledge of these matters is admittedly very limited. It does not seem likely, given the present state of knowledge and paltry research effort into moral enhancement, that we shall be able to effect any noticeable improvement in moral character and behaviour before it is very possible for some morally warped individuals misuse our potent scientific knowledge and technology with fateful consequences. To repeat, this is so because for this to happen, it may be enough that some tiny minority of humanity is morally corrupt.

Some may want to object that sufficient cognitive enhancement by itself will produce the moral enhancement required to avoid the misuses of science and technology we have indicated. They might believe this because they believe that people engage in highly immoral behaviour largely because they are enthralled by some religion or some other type of unfounded ideology which will disappear in the wake of cognitive enlightenment. We are, however, strongly inclined to doubt both that all highly immoral behaviour is always ideologically fuelled and that cognitive enlightenment will weed out all ideologies.

As regards the first point, pure selfishness could surely drive people to highly immoral behaviour. Consider, for instance, one of our biggest moral problems: the depletion of the planet's resources, which endangers the future of humanity (and other sentient beings). It seems not to be primarily due to religion or any other ideology, but to selfish desires for comfort, entertainment etc. which disregard the interests of future generations. As regards the second point, remember that some prominent scientists are fervent and fundamentalist religious believers.

We do not deny that cognitive enhancement is indispensable for moral enhancement. It could assist us in finding out what the moral truth is and which moral beliefs are justified. Theoretical rationality is important to practical rationality⁴⁶. But finding out what we are

justified in believing to be morally right does not automatically make us act accordingly. As already remarked, Socrates was overly optimistic in thinking otherwise.

5. CONCLUDING REMARKS

The picture we have painted is rather gloomy: given that we are unlikely to arrest the progress of cognitive enhancement, or that we may already be too smart for our own good, we are in need of a rapid moral enhancement, but such an enhancement could only be effected if significant scientific advances be made. In order to appreciate fully the problem we are raising, two general points should be emphasized. First, it is, as a rule, much easier to harm than to benefit. It is quite easy for virtually anyone to do serious harm – consider, for instance, the recent shooting at Virginia Tech on April 16, 2007, when Seung-Hui Cho killed 32 people in the worst civil shooting in US history. Cho used two semiautomatic handguns. The actual killings took place in a matter of minutes. It is almost never possible to save 32 lives in the same period of time. People can be killed at any point in their lives, but it is only in exceptional circumstances, such as when we can save them from death, that we can benefit them as much as we harm them when we kill them.

The comparative easiness of doing harm holds not only the individual level, but also on the macro-levels. For instance, it is much more difficult to improve significantly upon a comparatively well-ordered ecological system such as the one on Earth than it is to seriously damage it. It may well be forever impossible for us to improve upon this system to the same degree as we are now damaging it. This is a reason why the greatest caution is called for when we contemplate large-scale interventions with our ecological system. It is also a reason why the misuse of scientific discoveries by few isolated evil-doers poses such an alarming threat.

The second point is that it appears rational to reject an intervention which is very likely to produce some improvement if there is also *some* risk of it causing a grave enough harm, even

if the expected utility of the intervention is mathematically greater than its expected disutility. To give a simple illustration, suppose that your level of welfare could be assigned a number, say, 100 units. Suppose also that probabilities could be assigned numerical values. Consider now your choice of whether or not to participate in a lottery in which the probability that you will gain 2 units is 99%, but the risk that you will lose all of your 100 units is 1%. Then the expected value of the possible gain is nearly double the expected disvalue of the possible loss, and standard decision theory would declare rational participation in the lottery. But many of us would regard it as more reasonable not to participate. It would strike us as almost crazy to run even such a small risk as one in a hundred of losing everything in order to gain an increase in well-being which is relatively insignificant. Now, this may be our situation with respect to further scientific advances. These advances will almost certainly lead to a small increase of our already high quality of life — better food and lodging, better means of transport and communication, etc. — but at the cost of marginally increasing the risk of death in the near future, through the misuse of some these advances.

Our reasoning could be summarized by the following main claims:

1. It is comparatively easy to cause great harm, much easier than to benefit to the same extent.
2. With the progress of science, which would be speeded up by cognitive enhancement, it becomes increasingly possible for small groups of people, or even single individuals, to cause great harms to millions of people, e.g., by means of nuclear or biological weapons of mass destruction.
3. Even if only a tiny fraction of humanity is immoral enough to want to cause large-scale harm by weapons of mass destruction in their possession, there are bound to be some such people in a huge human population, as on Earth, unless humanity is extensively morally enhanced. (Or the human population is drastically reduced, or

there is mass genetic screening and selection, though we take it that there is no morally acceptable way of achieving these sufficiently effectively.)

4. A moral enhancement of the magnitude required to ensure that this will not happen is not scientifically possible at present and is not likely to be possible in the near future.
5. Therefore, the progress of science is in one respect for the worse by making likelier the misuse of ever more effective weapons of mass destruction, and this badness is increased if scientific progress is speeded up by cognitive enhancement, until effective means of moral enhancement are found and applied.

True, there are also respects in which scientific progress accelerated by cognitive enhancement would be for the better, by better protecting us against threats posed by asteroids, epidemics, etc. We have not attempted to settle definitely the balance between these good and bad respects. This would be exceedingly difficult, for it would involve accurate predictions of what this scientific progress will bring along, as well as assessments of the probabilities of different uses and misuses of these scientific discoveries.

At the very least, the perils of cognitive enhancement require a vigorous research program on understanding the biological underpinnings of moral behaviour. As Hawking quipped, our future may depend on making ourselves wiser and less aggressive. If safe moral enhancements are ever developed, there are strong reasons to believe that their use should be obligatory, like education or fluoride in the water, since those who should take them are least likely to be inclined to use them. That is, safe, effective moral enhancement would be compulsory.

Martial arts have for centuries given people superior powers of fighting. But it has been an integral part of many of the oriental martial arts to *morally* educate the combatant in the deployment of this power. They have been linked to spiritual development, which is seen as a prerequisite of learning the physical art. In the same way, moral enhancement should

accompany cognitive enhancement, since the latter is a means that could be put to both good and bad uses.

CS Lewis in *The Magician's Nephew*, describes the “Deplorable Word.” This is a “magical curse which ends all life in the world except that of the one who speaks it. Lewis has the lion Aslan say to the central characters, who are children from the Victorian era:

“ It is not certain that some wicked one of your race will not find out a secret as evil as the Deplorable Word and use it to destroy all living things. And soon, very soon, before you are an old man and an old woman, great nations of your world will be ruled by tyrants who care no more for joy and justice and mercy than the Empress Jadis. Let your world beware. That is the warning.”⁴⁷ ”

If we all knew the Deplorable Word, the world would likely not last long. The Deplorable Word may arrive soon, in the form of nanotechnology or biotechnology. Perhaps the only solution is to engineer ourselves so we can never utter it, or never want to utter it.

NOTES

1 Glover, J. (1999) *Humanity: A Moral History of the Twentieth Century* (London, Jonathan Cape)

2 Though there has been a steady increase in IQ over the last century, known as the Flynn effect, the significance of this is unclear.

3 Thanks to Anders Sandberg for research assistance. For a complete review of cognitive enhancement, see www.enhanceproject.org

4 Tang, Y. P., Shimizu, E., Dube, G. R., Rampon, C., Kerchner, G. A., Zhuo, M., Liu, G., Tsien, J. Z. (1999) ‘Genetic enhancement of learning and memory in mice’, *Nature* 401, 63-69.

5 Falls, W. A., Miserendino, M. J. D. & Davis, M. (1992) ‘Extinction of fear-potentiated startle: blockade by infusion of an NMDA antagonist into the amygdala’ *Journal of Neuroscience*, 12, 854-863.

-
- 6 Wei, F, Wang, G-D, Kerchner, G. A., Kim, S. J., Xu, H-M., Chen, Z-F., Zhuo, M. (2001) 'Genetic enhancement of inflammatory pain by forebrain NR2B overexpression', *Nature Neuroscience*, 4, 164 – 169.
- 7 Routtenberg, A., Cantalalops, I., Zaffuto, S., Serrano, P., Namgung, U., (2000) 'Enhanced learning after genetic overexpression of a brain growth protein', *Proc National Acadademy of Science U S A*. Jun 20; 97, 13, pp.7657-62.
- 8 Wang, H, Ferguson, G. D., Pineda, V. V., Cundiff, P. E., Storm, D. R., (2004) 'Overexpression of type-1 adenylyl cyclase in mouse forebrain enhances recognition memory and LTP', *Nat Neurosci*. Jun, 7, 6, pp.635-42.
- 9 Lee, E. H., Ma, Y. L. (1995) 'Amphetamine enhances memory retention and facilitates norepinephrine release from the hippocampus in rats', *Brain Research Bulletin*, 37,4, pp. 411-6. Soetens, E., D'Hooge, R., Hueting, J.E. (1993) 'Amphetamine enhances human-memory consolidation', *Neuroscience Letters*, Oct 14, 161, 1, pp. 9-12; Soetens, E., Casaer, S., D'Hooge, R., Hueting, J.E., (1995) 'Effect of amphetamine on long-term retention of verbal material', *Psychopharmacology (Berl)*, May, 11, 2, pp. 155-62.
- 10 Korol, D. L., Gold, P. E., (1998) 'Glucose, memory, and aging', *American Journal of Clinical Nutrition*, Vol 67, 764S-771S; Foster, J. K., Lidder, P.G., Sünram, S. I. (1998) 'Glucose and memory: fractionation of enhancement effects' *Psychopharmacology*, 137 : 259-270; Meikle, A., Riby, L. M., Stollery, B. (2005) 'Memory processing and the glucose facilitation effect: the effects of stimulus difficulty and memory load', *Nutritional Neuroscience*, Aug;8(4):227-32; Winder, R., Borrill, J., (1998) 'Fuels for memory: the role of oxygen and glucose in memory enhancement', *Psychopharmacology (Berl)*, Apr;136(4):349-56.
- 11 Gulpinar, M.A., Yegen, B.C., (2004) 'The physiology of learning and memory: role of peptides and stress', *Current Protein and Peptide Science*, Dec;5(6):457-73.
- 12 Iversen, S. D., (1998) 'The pharmacology of memory', *Comptes Rendus de l Academie des Sciences*, Serie III., Feb-Mar;321(2-3):209-15. Power, A. E., Vazdarjanova, A., McGaugh, J.L., (2003) 'Muscarinic cholinergic influences in memory consolidation', *Neurobiology of Learning and Memory*, Nov;80(3):178-93. Freo, U., Ricciardi, E., Pietrini, P., Schapiro, M.B., Rapoport, S.I., Furey, M.L.(2005) 'Pharmacological modulation of prefrontal cortical activity during a working memory task in young and older humans: a PET study with physostigmine', *American Journal of Psychiatry*, Nov;162(11):2061-70.
- 13 Mondadori, C. (1996) 'Nootropics: Preclinical Results in the Light of Clinical Effects; Comparision with Tacrine', *Critical Reviews in Neurobiology*, 10:3/4, 357-370
- 14 Lynch, G., (1998) 'Memory and the brain: unexpected chemistries and a new pharmacology', *Neurobiology of Learning and Memory*, Jul-Sep;70(1-2):82-100. Ingvar, M., Ambros-Ingerson, J., Davis, M., Granger, R., Kessler, M., Rogers, G.A., Schehr, R.S., Lynch, G.,(1997) 'Enhancement by an ampakine of memory encoding in humans', *Experimental Neurology*, Aug;146(2):553-9.
- 15 Lynch, G., (2002) 'Memory enhancement: the search for mechanism-based drugs', *Nature Neuroscience*, November Volume 5 Supplement pp 1035 – 1038
- 16 Meck, W. H., Smith, R. A., Williams, C.L. (1987) 'Pre- , Postnatal Choline Supplementation Produces Long-term Facilitation of Spatial Memory', *Developmental Psychobiology*, 21: 4, 339-353. Mellott, T. J., Williams, C.L., Meck, W.H., Blusztajn, J.K.(2004) 'Prenatal choline supplementation advances hippocampal development and enhances MAPK and CREB activation', *FASEB J*, Mar;18(3):545-7

-
- 17 Walsh, R., Budtz-Olsen, N., Penny, J. E., Cummins, R. A. (1969) 'The effects of environmental complexity on the histology of the rat hippocampus', *The Journal of Comparative Neurology*, Volume 137, Issue 3, Pages 361 – 365. Greenough, W. T., Volkmar, F. R., (1973) 'Pattern of dendritic branching in occipital cortex of rats reared in complex environments', *Experimental Neurology*, Volume 40, Issue 2, August, 491-504. Diamond, M. C., Johnson, R. E., Ingham, C.A., (1975) 'Morphological changes in the young, adult and aging rat cerebral cortex, hippocampus, and diencephalon', *Behavioral Biology* Jun;14(2):163-74. Nilsson, M., Perfilieva, E., Johansson, U., Orwar, O., Eriksson, P. S. (1999) 'Enriched environment increases neurogenesis in the adult rat dentate gyrus and improves spatial memory', *Journal of Neurobiology*, Jun 15;39(4):569-78.
- 18 Farlow, M., Anand, R., Messina, J. Jr, Hartman, R., Veatch, J. (2000) 'A 52-week study of the efficacy of rivastigmine in patients with mild to moderately severe Alzheimer's disease', *European Neurology*, 44(4):236-41.
- 19 Barch, D.M. (2004) 'Pharmacological manipulation of human working memory', *Psychopharmacology (Berl)*, Jun;174(1):126-35. Epub 2004 Jan 30.
- 20 Müller, U., Steffenhagen, N., Regenthal, R., Bublak, P. (2004) 'Effects of modafinil on working memory processes in humans', *Psychopharmacology*, 177: 161-169.
- 21 Turner, D. C., Robbins, T. W., Clark, L., Aron, A. R., Dowson, J., Sahakian, B. J. (2003) 'Cognitive enhancing effects of modafinil in healthy volunteers', *Psychopharmacology*, Volume 165, Number 3, January, 260 – 269.
- 22 See Sandberg, A. and Bostrom, N. (2006) 'Cognitive Enhancement: A Review of Technology', *ENHANCE*
<http://www.enhanceproject.org/> and <http://www.enhanceproject/enhance/> for a comprehensive review.
- 23 Muraven, M., Baumeister, R. F. (2000) 'Self-Regulation and Depletion of Limited Resources: Does Self-Control Resemble a Muscle?' *Psychological Bulletin* 2000, Vol. 126, No. 2, 247-259 .
- 24 Baumeister, R. F., Bratslavsky, E., Muraven, M., Tice, D. M. (1998) 'Ego Depletion: Is the Active Self a Limited Resource?' *Journal of Personality and Social Psychology*, Vol. 74, No. 5, 1252-1265.
- 25 Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., & Plant, E. A. (2006). 'Self-control relies on glucose as a limited energy source: Willpower is more than a metaphor.' Submitted for publication.
<http://www.psy.fsu.edu/~baumeistertice/gailliotetal2006.doc>
- 26 Lykken, D. T., (2005) 'Mental energy', *Intelligence*, Volume 33, Issue 4, July-August, 331-335.
- 27 Jamison, K. R., (2004) *Exuberance*, (New York: Alfred A. Knopf)
- 28 Teitelman, E. (2001) 'Off-label uses of modafinil', *American Journal of Psychiatry* 158, 1341. Myrick, H., Malcolm, R., Taylor, B., LaRowe, S., (2004) 'Modafinil: preclinical, clinical, and post-marketing surveillance--a review of abuse liability issues', *Annals of Clinical Psychiatry*. Apr-Jun;16(2):101-9. Buguet, A., Moroz, D.E., Radomski, M.W. (2003) 'Modafinil--medical considerations for use in sustained operations', *Aviation, Space, and Environmental Medicine*, Jun;74(6 Pt 1):659-63.
- 29 Caldwell, J. A. (2001) 'Efficacy of stimulants for fatigue management: The effects of Provigil and Dexedrine on sleep-deprived aviators', *Transportation Research Part F: Traffic Psychology and Behaviour*, Part F:19-37.

30 Batejat, D.M., Lagarde, D.P., (1999) 'Naps and modafinil as countermeasures for the effects of sleep deprivation on cognitive performance', *Aviation, Space, and Environmental Medicine*, May;70(5):493-8.

31 Gill, M., Haerich, P., Westcott, K., Godenick, K. L., Tucker, J. A. (2006) 'Cognitive Performance Following Modafinil versus Placebo in Sleep-deprived Emergency Physicians: A Double-blind Randomized Crossover Study', *Academic Emergency Medicine*, Volume 13, Number 2 158-165

32 Rees, M. (2003) *Our Final Century* (London: William Heinemann). p. 58

33News 3 (2006) *The Guardian*, Thurs Aug 3,

34 Kurzban, R, Tooby, J, and Cosmides, L.(2001) 'Can race be erased? Coalitional computation and social categorization' *Proceedings of the National Academy of Sciences* 98:15387-15392. These authors argue exposure to alternate social worlds reduces the tendency to categorise by race, showing there may be important social means of making us more moral.

35 See e.g. Wright, R (1994) *The Moral Animal*, (New York: Pantheon) chap. 9.

36 Thanks to Wasserman for directing our attention to this point

37 Sober, E., Wilson D. S. (1998) *Unto Others*, (Cambridge MA: Harvard U. P.

38 de Waal, F., Eds. Macedo, S., Ober, J. (2006) *Primates and Philosophers*, (Princeton: Princeton U. P)., p. 43

39 de Waal, F., Eds. Macedo, S., Ober, J. (2006) *Primates and Philosophers*, (Princeton: Princeton U. P), pp. 47-8

40 de Waal, F., Eds. Macedo, S., Ober, J. (2006) *Primates and Philosophers*, (Princeton: Princeton U. P)p. 48

41 de Waal, F., Eds. Macedo, S., Ober, J. (2006) *Primates and Philosophers*, (Princeton: Princeton U. P)p. 49

42 K. Jensen, J. Call and M. Tomasello (2007) 'Chimpanzees Are Rational Maximizers in an Ultimatum Game', *Science*, 318, 5 October, 2007, pp.107-9

43 B. Wallace, D. Cesarini, P. Lichtenstein and M. Johannesson (2007) 'Heritability of Ultimatum Game Responder Behavior', *Proceedings of the National Academy of Sciences*, 104, 40, 15631-4

44 Savulescu J., Momeyer R.W. (1997) Should Informed Consent Be Based on Rational Beliefs? *Journal of Medical Ethics*;23, 5, pp. :282-8

45 It is also possible, of course, that as these and other conditions are understood, interventions are developed which seriously harm humans and human society, such as by promoting docility, blind subordination to authority and loss of curiosity. This represents the misuse of technology which itself is a manifestation of the problem presented by the rapid advance of technology, though at a higher order. We may require moral enhancements at the highest order in order to prevent the modification of our dispositions to accept or realize very bad states of the world.

46 Op. Cit

47 'The Deplorable Word', *Wikipedia*, http://en.wikipedia.org/wiki/Deplorable_word accessed 15 January 2007 Thanks to Toby Ord for this example.