



Unfit for the Future: The Need for Moral Enhancement

Ingmar Persson and Julian Savulescu

Print publication date: 2012

Print ISBN-13: 9780199653645

Published to Oxford Scholarship Online: September 2012

DOI: 10.1093/acprof:oso/9780199653645.001.0001

Moral Enhancement as a Possible Way Out

Ingmar Persson

Julian Savulescu

DOI:10.1093/acprof:oso/9780199653645.003.0010

Abstract and Keywords

To enable liberal democracies to stop human induced climatic and environmental deterioration it is then necessary to morally enhance their citizens, to make them appropriately concerned about the interests of future generations and of people world-wide. Since traditional methods of moral enhancement hitherto have been insufficient, this chapter suggests that we should also explore biomedical means of moral enhancement. It is argued that the central moral dispositions of altruism and a sense of justice have biological bases that are sensitive to drugs such as oxytocin. It is also argued that moral enhancement by biomedical means undercuts neither human freedom and responsibility, nor the influence of moral reasoning. However, it is admitted that research into biomedical agents of moral enhancement is still in its infancy and that it is too early to judge its fruitfulness. There is also the general difficulty that means of moral enhancement have to be administered by morally imperfect people.

Keywords: altruism, biomedical means, freedom, justice, moral enhancement, moral reasoning, responsibility, oxytocin

If we look back upon the 80,000 years or so that have passed since *Homo sapiens* began to colonize the Earth from Eastern Africa, we discern a process of relentless expansion and exploitation, with very few episodes of restraint. When there has been restraint, the circumstances have been significantly different from those that are now confronting us globally. Societies have been able to adopt sustainable policies, which enabled them and their descendants to

overcome ecological threats without colonization or territorial expansion when either (1) these societies were so small that everyone knew everyone else, so that mutual concern and trust were possible, or (2) they featured a wise dictatorial power, which ruled over all their territory and their inhabitants (Diamond, 2005: 277–8, 429).

Obviously, neither of these conditions holds with respect to the resources of the Earth. Here it is not only the case that there is a multitude of states competing with each other, there is also a gross inequality in respect of welfare between these nations, with nations in the developing world aspiring to the wealth that Western democracies have enjoyed for so long. It seems unlikely that these deep conflicts could be bridged in the short time at our disposal before we have damaged the global climate and environment irreparably. Therefore, the most likely course of events might well be that we shall walk into some sort of global collapse, the magnitude of which is hard to divine. This is confirmed by the fact that, though climatic and environmental problems have been widely discussed for a couple of decades, and with a special intensity in the last few years, precious little has been done in practice to mitigate appreciably the detrimental anthropogenic impact upon the global climate and environment. This is shown, for example, by the so-called ‘Overshoot Day’, the day when we have used as much resources as the Earth can regenerate in a **(p.101)** year and released as much waste as it can reabsorb in a year. Overshoot day has steadily occurred earlier and earlier in the last couple of decades, and has not recently been pushed back from its alarmingly early occurrence in the year. Since 2008 it is estimated to have occurred in August or September (see *Global Footprint Network*). That is, in a year humans spend close to 30 per cent more than that which the Earth can provide in the same period of time.

There is reason to fear, then, that humankind will follow the fate of previous, now extinct human cultures, like the famous one on Easter Island (Rapa Nui), which have brought about their own downfall by overexploitation of natural resources—a human parallel to the non-human animals, which have been so reproductively successful that they have caused their own populations to collapse. In addition, we must take into consideration a scenario in which shrinking natural resources like oil, arable land, and water may provoke wars with weapons of mass destruction.

Nonetheless, we should not regard a global collapse as inevitable. We are not biologically or genetically determined to go on consuming voraciously, until we are forcibly stopped by the depletion of natural resources. As our history eloquently shows, we are more than any other animal biologically or genetically disposed to learn by experience, and we are now learning that our present course of action spells disaster. We *can* decide to overturn any predictions made

about what we will intentionally do because no prediction can take into account the effects that it itself will have.¹

But, although our behaviour is highly adaptable in the light of experience, some of our behavioural patterns are propelled by quite recalcitrant drives. For instance, it is practically certain that, if the human species does not undergo anything like a dramatic genetic mutation, a majority of humans will not of their own free starve themselves, abstain from sex, and seek utter solitude, whatever their experience seems to teach them. True, some eccentrics might adopt these unusual forms of behaviour, but we can rest assured that they will not spread to the majority of humans. Consequently, we can in practice exclude the possibility of future societies in which these forms of behaviour are the rule.

(p.102) The question now is whether the motivational drives that prod humans to overexploit their environment are too recalcitrant in the majority of them to be hemmed in by an insight into how destructive this conduct is. It is likely that this insight could effect such a behavioural modification in a minority. These people might also be given training in science, which would render them better able to understand the mechanisms behind such phenomena as human induced climate change. This is why it would be possible that there be a meritocracy intent on averting a serious environmental breakdown. But in order for a democracy to avert it, the majority has to be converted. We have seen reasons to believe that this would not be easily achieved, although it is certainly possible because human beings are to a great extent malleable by the norms in the society in which they grow up, and radical social revolutions are conceivable.

Societies have always taken advantage of this fact of human malleability by imprinting upon their subjects moral norms conducive to the survival and prosperity of these societies. The norms inculcated have included not merely norms about refraining from certain types of behaviour, such as killing innocent members of one's own society against their will, or stealing their property, but also norms about making positive contributions to the public good, by helping needy members of one's own community and defending the community against external enemies. Today, when scientific technology has vastly increased our powers of action, and has connected societies all over the world with each other by means of travel and commerce, such intra-societal norms are not nearly enough. Liberal democracies need to inculcate norms that are conducive to the survival and prosperity of a world-community of which their societies are integral parts. They have to take the step from a social liberalism, which acknowledges the need for state interference to neutralize the glaring welfare inequalities within a society, to a *global(ly)* responsible liberalism, which extends welfare concerns globally and into the remote future. This is something that liberal democracies have largely failed to do so far, but it seems necessary that

they do so straightaway if they are not to undermine themselves and the living conditions on Earth.²

(p.103) This development of liberalism is called for because, as we have repeatedly stressed, owing to the progress of science, the range of our powers of action has widely outgrown the range of our spontaneous moral attitudes, and created a dangerous mismatch. Environmental problems arise because there is a growing domain, which we could affect by our actions but which is in the periphery of our moral consciousness. To come to grips with these problems we believe that it is necessary to widen the horizons of our moral consciousness. It strikes us as wishful thinking to believe that these problems can be wholly solved by technological innovations like clean fuel and carbon sequestration (contrast Posner, 2004: 160). Even if the carbon emission per unit of GDP produced is cut by as much as 90 per cent until 2050 that would be far from enough to keep the temperature rise at a safe level (Hamilton, 2010: 46). As we have already remarked, human beings have a well-attested propensity to be too optimistic or over-confident about their own abilities and about what they can achieve in the future.

Consider the problem of giving aid to the developing world, discussed in Chapter 4. Here we have had in our possession for a long time the means to eliminate a lot of the starvation and diseases in poor countries, but the will has been lacking to apply these means fully. Likewise, various kinds of cleaner technology exist today—perhaps to such an extent that they could cater for our entire need for energy—but they are not completely implemented. For instance, it is technologically possible to use hydrogen-powered fuel cells as a non-polluting energy source for cars. But there is little demand for such cars because they are much more expensive than internal combustion engine cars powered by petrol, and there are virtually no hydrogen fuelling stations. However, the price on fuel cell cars is not likely to go down, and a sufficient number of fuelling stations are not likely to be built, until the demand for these cars increases. There is, then, something of a catch 22 here.

James Martin speculates, in line with our reasoning in Chapter 8, that because of its ‘all-powerful central government’, ‘low-cost manufacturing’, and ‘excellent research,’ China could quickly start to ‘mass-produce and export ecologically benign cars in vast numbers’, and thus make an **(p.104)** enormous profit out of a potentially growing demand for this and other environmental-friendly products (2006: 22–3). But there is an initial obstacle: the protracted transition phase in which the production of such cars will not be profitable because there is not the infrastructure to make them attractive to buyers. So, the Chinese government will probably not have an incentive to change the course of their domestic production. Indeed, this seems to be precisely what is happening as the Chinese government provides petrol at lower cost than the market price to stimulate an automobile revolution of the nation. Another reasonable fear is that instead of

developing clean energy, it will fall back upon the enormous coal reserves that China sits upon, though it is comparatively dirty coal. Still, as we have already observed, the future political course of China seems to be harder to predict than that of the US.

Our conclusion is, then, that the solution to climatic and environmental problems is not wholly technological. Nor will there be a political solution in the democratic form of government, unless the will to act morally grows stronger in the public. For these problems have to do with such matters as people being too little concerned about others who are beyond their immediate circle of acquaintances, especially large numbers of such strangers, too much preoccupied with the present and imminent future, and feeling too little responsible for their omissions and collective contributions. Without a willingness to make personal sacrifices for the sake of people in remote countries and in the remote future, there will in all probability not be enough of an effort to develop and put to full use a technology that could arrest or significantly lessen anthropogenic climatic and environmental degradation. To develop this technology requires financial resources that must be extracted in ways that impose costs upon us, like taxes or restrictions on fossil fuels and on certain kinds of food, and to put this technology to use requires our willingness to adopt more costly alternatives to present practices.

Liberal democracies have responded to moral problems in the current global setting by supporting the doctrine of the equality of all humans, e.g. via international organizations like the UN. This egalitarianism is major step forward in comparison to the racism and sexism that were still prevalent in a not very distant past,³ but this egalitarian ideology has a **(p.105)** long way to go before it succeeds in stamping out the deep-seated xenophobia of our nature. It is still pretty much of a façade covering an underlying xenophobia which under certain circumstances, e.g. when resources become scarce, breaks out in excessive violence. For instance, this happened in ex-Yugoslavia and Rwanda in the 1990s, and it has more recently happened in Dafur. It is not unreasonable to fear that when crucial natural resources, like oil or water, become scarce in the imminent future, this will cause an outburst of wars with weapons of mass destruction between nations that are 'foreign' or 'out-groups' to each other, racially or religiously.

Moreover, the egalitarian ideology has not pervaded the fundamental domain of economy. Here a Lockean theory of property rights continues to rule the day, even though Locke's own proviso that property acquisition is legitimate only as long as 'there is enough, and as good left in common for others' (1690/1990: II. v. 27) ceased to hold a long time ago. In terms of difference in per capita income between the richest and poorest countries, the world now appears more unequal than it has ever been: 'the difference between the per capita incomes of the richest and the poorest countries was 3 to 1 in 1820, 11 to 1 in 1913, 35 to 1 in

1950, 44 to 1 in 1973, and 72 to 1 in 1992' (Seitz, 2008: 3). At the beginning of the present millennium the wealthiest fifth of the world's population stood for 86 per cent of the world's GDP, while the poorest fifth stood for only 1 per cent of it, and the richest three individuals owned as much as did 600 million people in the poorest countries. Economic inequality in affluent societies, even those with a strong egalitarian ideology, like the Scandinavian countries, increases rather than decreases: for instance, in Sweden the gap between the average salary of a CEO and an average factory worker is now as big as it was in the 1950s after having decreased in the 1960s and 1970s. The tendency in the US is even scarier: the ratio rose from 42:1 in 1960 to a staggering 531:1 in 2000.

To cope with climatic and environmental problems, as well as the problem of global inequality, the ideology of human equality must exercise a stronger motivational influence and overcome the limitations of our altruism and sense of justice. But, to repeat, we must also overcome the **(p.106)** bias towards the near future, our numbness to the suffering of great numbers, and our weak sense of responsibility for our omissions and collective contributions. It should be asked to what extent this moral enhancement could be accomplished by traditional methods of moral education. These methods include such things as carefully reflecting on the reasons for which actions are morally right or wrong, and making as vivid to oneself as possible how one's actions affect others. This could be made vivid by regularly imagining, actually confronting, or watching films of the suffering victims of wrongdoing. But a ground for suspecting that by such measures moral enhancement could not be accomplished to a sufficient degree in time to avert disastrous misuses of modern technology is that the degree of moral improvement in the 2,500 years that have elapsed since the first great teachers of morality appeared is nowhere near matching the degree of technological progress during the same period.

To bring out this point, it might be helpful to distinguish between improvement in respect of (1) moral doctrines and (2) moral actions and reactions, which requires that improvements in respect of moral doctrine are internalized to the degree that they regulate conduct. We have noted that there have been improvements in respect of moral doctrine to the effect that, for example, racial and sexual differences are now widely regarded as morally irrelevant. It could also be mentioned that punishments in most societies, have become a great deal less barbarous in recent times. Nonetheless, it seems that this improvement has been modest in comparison to the formidable improvement as regards scientific and technological knowledge. By itself this doctrinal difference might not be important, but it assumes greater significance when we add that this scientific and technological knowledge can extend our powers of action by being used to build ever more sophisticated machines and other devices, which can be mastered after a relatively short period of training. In contrast, it is quite hard to internalize moral doctrines to the degree that they determine our behaviour. This is shown not only by how few people live up to more demanding doctrines,

for example, those that require the sacrifice of a substantial part of their welfare to save the life of strangers, but more emphatically by the frightening speed with which people, when political conditions allow it, are capable of regressing to barbarous behaviour, which one had hoped humanity had left behind for good. Every new generation has to go through a strenuous moral training anew. Consequently, there is a widening gap between what we are practically **(p.107)** able to do, thanks to modern technology, and what we are morally capable of doing, though we might be somewhat more morally capable than our ancestors were. It is this motivational internalization of moral doctrines that we think could be sped up by means which the scientific exploration of the genetic and neurobiological bases of our behaviour might put into our hands. We call moral enhancement by such means *moral bioenhancement*; possible examples of moral bioenhancement would be drug treatment and genetic engineering.⁴

Steven Pinker hypothesizes that ‘enhanced powers of reason—specifically, the ability to set aside immediate experience, detach oneself from a parochial vantage point, and frame one’s ideas in abstract, universal terms—would lead to better moral commitments’ (2011: 656). He believes that the requisite enhanced powers of reason are evidenced by the so-called Flynn Effect, the fact that in the 20th century the IQ has gone up by an average of three points per decade (2011: 650–1). His hypothesis is that this rise could explain ‘the documented declines of violence in the second half of the 20th century’ (2011: 656): the relatively scarcity of bigger wars and genocide; the marked drop of homicide rates and rates of other violent crimes against women, ethnic minorities, homosexuals, and children; more humane forms of punishment; the recognition of the equal rights of all people irrespective of race and sex; less cruelty to non-human animals; etc.

Now, we do not want to deny that enhanced powers of reason are tremendously important for moral enhancement—perhaps Pinker is right that they are the main force behind the moral improvements that he lists. But we do want to deny that once reason ‘is programmed with a basic self-interest and an ability to communicate with others, its own logic will impel it, in the fullness of time, to respect the interests of ever-increasing numbers of others’ (2011: 669). We do not see how such an expanding circle of concern is possible without the assistance of the moral dispositions of altruism and a sense of justice. Reason and self-interest could surely tell you to rob and kill an injured stranger in the wilderness rather than help him, or to abstain from returning a favour to someone you will not ever see again rather than to return it at a cost to yourself.

(p.108) Pinker suggests that our sympathy or compassion with strangers is too feeble to prod us to make any costlier sacrifices to help strangers even when millions of them are in need of help. True, it is feeble, and there is no way it could be made proportionate to the suffering of millions. But if we had not felt *any* sympathy, we could have paid as little attention to the suffering of millions of people as to the millions of grains of sand stirred up by a gust of wind because

neither affects our self-interest. If, however, our sympathy is aroused by the sufferers, we will pause and reflect upon their plight. We will then realize in more detail how horribly great the amount of their suffering is, and our sympathy will receive a boost that could animate us to help them at our own expense. Moreover, if reason informs us that the racial or other differences between one group of individuals for whom we feel sympathy and a sense of justice and another group for which we do not possess these attitudes are insignificant, our sympathy and sense of justice will spread from the former group to the latter. But there can be no such spreading if these attitudes are lacking all together. Therefore, we think that sympathy and a sense of justice are indispensable for being fully moral, and that the explanation of why humanity so far has failed to deal with climate change and environmental destruction—in spite of the enhanced powers of reason—is that they leave self-interest untouched and call upon our insufficient sympathy and sense of justice as regards future generations and non-human animals.

As we see it, then, the core moral dispositions, which are the foremost objects of moral enhancement, are altruism and a sense of justice as it primarily manifests itself in tit-for-tat. By classifying these as moral dispositions, we imply that, by themselves, they *always* issue in a morally correct treatment of the individuals to whom they are directed. To be sure, if you are more strongly altruistically disposed towards some individuals than others, this might result in your giving an unfair advantage to the former individuals, but *taken by itself* the behaviour that the strong altruist disposition towards one individual issues in is not morally wrong.⁵ In addition, we have also surveyed some cognitive dispositions which are **(p.109)** morally relevant dispositions rather than moral dispositions, like the bias towards the near future and the conception of responsibility as being causally-based. They are morally relevant because they limit the moral dispositions of altruism and justice, but their scope is wider, comprising, for example, self-regarding prudence. What we mean by moral enhancement is first and foremost enhancement of the latter two central moral dispositions, but since the reduction of the bias towards the near future and the conception of responsibility as being causally-based could result in extensions of altruism or the sense of justice, moral enhancement in a wider sense encompasses the reduction of both the temporal bias and the commitment to a causally-based conception of responsibility.

We have suggested that altruism and the sense of justice have biological bases by sketching their evolutionary origin (this does not exclude that they can be influenced also in significant ways by cultural or other social means). But this hypothesis of a biological basis is also supported by studies of animals and identical twins. With respect to animal studies, it is important to make explicit that what we mean by altruism is something else than mere emotional contagion, for example, fear spreading through a herd. It is also something else than the distress and helping behaviour elicited by nothing but the *outward*

signs of another's suffering, perhaps for the reason that one finds these signs unpleasant.⁶ We have in mind a more sophisticated reaction which involves (1) *empathy*, i.e. a capacity to imagine what it would be like to be another conscious subject and feel its pleasure or pain, etc. as well as (2) *sympathetic concern* about the well-being of this subject for its own sake, e.g. an intrinsic desire to relieve pain, occasioned by the empathic act of imagination. No doubt, it is difficult to tell whether a non-human animal is capable of altruism in this sense, but it seems likely that this is so if an animal exhibits helping behaviour tailored to the individual needs of another animal 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). Now there is reason to believe that, alongside humans, at least apes and dolphins, and perhaps elephants are capable of altruism in this sophisticated sense. Chimpanzees have often been observed performing acts of low-cost altruism to strangers and acts of **(p.110)** high-cost altruism to those who are near and dear (de Waal, 2010: ch. 4). Evidence of dolphin and elephant altruism is less extensive, but existent (de Waal, 2010: 125–39).

The occurrence of the tit-for-tat strategy in animals 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' (2006: 43; 2010: 173–4). This looks suspiciously like gratitude. In another of his studies, capuchin monkeys (which have the largest brains relative to body size of all monkeys) were paired with a group mate. 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' (2006: 47–8).

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' (2006: 48; cf. 2010: 187 ff.). He stresses that the reactions of the capuchin monkeys were rather 'egocentric' (2006: 49) in the sense that they reacted negatively only when they themselves were treated worse, not when their partners got the worse deal. Therefore, it may not be accurate to speak of a sense of fairness without qualification. A proper sense of fairness

might require an ability to apply the notion across the board and, so, perhaps presupposes an ability to empathize, as does genuine altruism. According to de Waal, there are indications that chimpanzees exhibit such a sense of fairness (2010: 190–3).

The hypothesis that the sense of justice has a biological basis has been confirmed by studies of human twins playing the roles of proposer and responder in the ultimatum game. Björn Wallace and associates have found that in the case of identical twins (who share the same genes), **(p.111)** 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 twins (2007: 15631–4). This indicates that the human sense of fairness has a genetic basis. According to Simon Baron-Cohen (2003: 114), there is also a striking correlation in respect of altruism in identical twins.

Furthermore, it is plausible to think that in general women have a greater capacity for altruism than men. If a general difference as regards this trait tracks gender, this is good evidence that the trait is biologically based. It has been argued at length by Baron-Cohen (2003) that as rule women have a greater capacity for ‘empathy’ than men. On our conception of empathy it is, as already remarked, a capacity to imagine vividly what it would be like to be another, to think, perceive, and feel as they do. Thus, as we conceive it, empathy does not involve any motivational component. On this conception, empathy is a merely component of altruism, as we understand it, since we take altruism to include also a motivational component of sympathetic concern about how others feel; a concern that they feel well rather than suffer. This is roughly how Baron-Cohen understands empathy⁷ and, so, his claims about empathy can be treated as equivalent to claims about altruism in our terminology.

Baron-Cohen notes that empathy can act as ‘brake on aggression’ (2003: 35). Thus, we should expect that a lesser male capacity for empathy is likely to go with the greater display of male aggression, which is borne out by the statistics of violent crimes like murder (Baron-Cohen, 2003: 36). Baron-Cohen does not maintain that women are not aggressive at all. His claim is rather that female aggression tends to take the subtler forms of backstabbing, social exclusion, etc. instead of direct physical assault, and these subtler forms of aggression presuppose mind-reading (2003: 35). He also reports that autism, which consists in a deficiency of at least the cognitive or imaginative aspect of empathy, is ten times more common among men (2003: 137). If this is right, it seems that in principle we could make men in general more moral by biomedical methods through making **(p.112)** them more like the men who are more like women in respect of sympathy and aggression, but without the tendency to social forms of aggression.⁸

Some critics of moral bioenhancement have feared that it would be corrosive of freedom and, thereby, of moral responsibility (e.g. Harris, 2011). But this example should make us realize that this fear is misguided: women are not less free and responsible than men because by biological nature they are more altruistic and less aggressive. Suppose, first, that our freedom and responsibility is compatible with it being fully causally determined whether or not we shall do what we take to be good. Then a judicious use of effective techniques of moral bioenhancement to increase a sense of justice and altruism will not reduce our freedom and responsibility; it will simply make it the case that we are more often, perhaps always, causally determined to do what we take to be good. It will do so by amplifying those biological factors that by nature are strong in those of us who are morally better. We would then act as a morally better person now acts. Such a person is not less free and responsible than those of us who more frequently fail to do what we think is morally right.

Suppose, on the other hand, that we are free and responsible only because, by nature, we are not fully causally determined to do what we take to be morally right. Then moral bioenhancement cannot be fully effective because its effectiveness is limited by the indeterministic freedom that we possess. So, irrespective of whether causal determinism or indeterminism reigns in the realm of human action, moral bioenhancement will not curtail human freedom and responsibility. Biomedical manipulation cannot change the basic laws of our behaviour by making us more (or less) causally determined; it simply uses knowledge of those laws to influence our behaviour.

However, some critics of moral bioenhancement seem to think that it would turn us into mindless robots who do not act for reasons. For instance, John Harris writes that moral bioenhancement will ‘make the freedom to do immoral things impossible, rather than simply make the doing of them wrong and giving us moral, legal and prudential reasons to refrain’ (2011: 105). But, in our view, those who have undergone moral **(p.113)** bioenhancement would act for the same reasons as those of us who are most moral today do, and the sense in which it is ‘impossible’ that they do what they regard as immoral will be the same for the morally enhanced as for the garden-variety virtuous person: it is something that it is psychologically or motivationally out of the question that they choose to do. We imagine that the moral motivation of those of us who are less morally motivated be increased so that it becomes as strong as the moral motivation of those of us who are by nature most morally motivated, not that this moral motivation be increased to the point at which it becomes *irresistible*, like a kleptomaniac’s desire to steal. The strength of the desire to do the morally right thing should be proportional to the reasons we have. It is a mistake to believe that people who are by nature morally good and always try to do what they regard as right are necessarily less free and responsible than those of us who more often fail to do so. Just as naturally virtuous people do not compulsively do

what they regard as right, so morally enhanced people will not compulsively do what they regard as right.

It is our view that some children should be subjected to moral bioenhancement, just as they are now subjected to traditional moral education. This is because the capacity to influence development under way is likely to be greater than the capacity to alter established motivational dispositions and behaviour. There is no reason to assume that moral bioenhancement to which children are exposed without their consent would restrict their freedom and responsibility more than the traditional moral education to which they are also exposed without their consent. It is of course true that if some children become more morally motivated through bioenhancement, they have not *chosen* to be more moral, but this is also true of the children who become more moral as the result of early moral education and natural endowments. It is quite unlikely that later in life the morally bioenhanced individuals will regret the fact they have undergone this treatment, since otherwise they might have been criminals who would have been punished and condemned by society. Thus, we cannot see that if children are exposed to moral bioenhancement, this will disrupt their freedom and responsibility more than when they are exposed to traditional moral education. As we have already noted, it is not true that they will compulsively do what they think is morally right; nor does moral bioenhancement exclude any options that they would have liked to be open.

It is, however, true that if we come into possession of very effective techniques of bioenhancement, and sharpen the use of traditional moral **(p.114)** education, we could determine what motivational states people will be in (if determinism rules in the realm of human behaviour). But this would not imply that these people are not responsible: people can be responsible for how they act and react in situations, even though someone else has determined how they will act and react in those situations. This is true even when the determination takes the form of coercion which restricts the subject's freedom. For instance, if someone forces you to hand over money at gunpoint, you are responsible for giving in to the threat (though this might be perfectly reasonable and not anything for which you are blameworthy). However, when determination amounts to moral enhancement of a person, it does not restrict freedom; it rather extends it, by making the subject more capable of overcoming urges which counteract the doing of what is seen as morally good. This is a point that should be emphasized: when we influence the motivational states of people, this could be liberating rather than constraining. It could be influence of a sort that they have reason to welcome rather than to eschew.

Harris's core claim about freedom, expressed in the idiom of Milton's *Paradise Lost*, seems to be that 'sufficiency to stand is worthless, literally morally bankrupt, without freedom to fall' (2011: 110). In other words, a decision to act in a way that is morally right is morally worthless—meaning, presumably, that

you are not morally praiseworthy for it—if you are not free not to make this decision, but instead a decision to do something wrong. This view can be shown to be mistaken by a kind of argument made famous by Harry Frankfurt (1969). Imagine that you decide to do the morally right thing on the basis of considering reasons for and against, as somebody who is morally responsible is supposed to do. Imagine, however, that there is a freaky mechanism in your brain which would have kicked in if you had been in the process of making, not this decision, but a decision to do something which is morally wrong. The mechanism would then irresistibly have made you decide to do the morally right thing. Hence, you are not free to fall, i.e. you cannot avoid deciding to do the morally right thing. Would the presence of this freaky mechanism then mean that you are not praiseworthy for making the right decision? It is hard to see why it would: after all, the mechanism was never called into operation; it remained idle. In fact, you decided to do the morally right thing for precisely the same reasons as someone whose brain does not feature the freaky mechanism could do, and whose praiseworthiness therefore is not in doubt. It seems plausible to think that what **(p.115)** determines whether you are morally responsible and praiseworthy is the *actual* occurrences that led up to your decision, not some merely hypothetical occurrences that could have led up to your decision, but in fact did not. Certainly, owing to the presence of the freaky mechanism, you are not free to decide to act immorally; this is not anything you could do. But, as already the example of handing over the money case at gunpoint brings out, freedom of will or action is not indispensable for moral responsibility. So, Harris's 'freedom to fall' is not essential for moral choice and action.

Harris also objects that moral bioenhancement could not be made to target the right sort of motivational states for 'the sorts of traits or dispositions that seem to lead to wickedness or immorality are also the very same ones required not only for virtue but for any sort of moral life at all' (2011: 104). However, a low level of altruism and high level of physical aggression, which lead to immoral behaviour, are not requisite for a moral life. To be sure, a certain amount of aggression might be necessary for moral behaviour: for instance, anger as a response to offences might be necessary to make offenders change their ways. But it should be a degree of anger which is proportionate to the size of the offence, and a sense of justice is required to ensure this, as well as that anger is not directed at people who do not deserve it.⁹

It is somewhat surprising to find Harris among the opponents of moral bioenhancement in view of the fact that he is in favour of cognitive enhancement by means of biomedical methods. A more common attitude is opposition to enhancement of all of our mental or psychological properties by such methods. Discussing cognitive or intellectual bioenhancement in particular, Nick Bostrom and Toby Ord (2006) hypothesize that a main cause of such resistance is a status quo bias, an irrational tendency to prefer a current state of affairs to a change of it simply because it already obtains. We are inclined to think that in many

instances the status quo bias is also part of the explanation of the opposition to moral bioenhancement, though this is presumably not so in Harris's case. But, in contrast to Bostrom and Ord, we do not regard the status quo bias as wholly irrational. This is because we link it with the fact that it is easier for us to harm than to benefit because changes to complex systems are more likely to be for the worse rather than for the better. Therefore, it is rational to be cautious **(p.116)** about changing a current state of affairs into something new, and the more radical the change, the more cautious it is rational to be because of the increased risk of unforeseen effects most of which are likely to be for the worse. By contrast, Bostrom and Ord claim that with respect to cognitive enhancements, uncertainty about the consequences 'far from being a sufficient ground for opposing them, is actually a strong consideration in their support' (2006: 669). For the reason already given, we disagree (cf. Agar, 2010: 138–9). On the other hand, it should be admitted that for most of us the status quo bias is too strong, so strong that it often makes us averse to changes for which we have very good reasons. This is especially likely when status quo has obtained for a longer period of time because we have a tendency to form an exceedingly strong emotional bond or attachment to things to which we have grown accustomed.

Since we have been employing the term 'empathy', it should again be underlined that we are here using it in an extended sense, such that empathy includes sympathy or a concern for the well-being of others, not merely imagining what the experiential state of another is like. Psychopaths are known for being skilful manipulators; perhaps this is because they have empathy in the narrower sense of a power to imagine from the inside what it would be like to be another. But they definitely lack sympathy and, so, are not altruists, or empathetic in the more inclusive sense (Baron-Cohen, 2003: 34). Sadists might also have empathy in the less inclusive sense—otherwise they could perhaps not fully enjoy the suffering of another—but they are certainly lacking in sympathy.

Even though it is true that our moral dispositions have a biological basis and, thus, are open in principle to manipulation by biomedical techniques, it does not follow that it is impossible to influence them by moral education. A rough analogy might be this: suppose that we want to improve someone's command of English. To a considerable extent, this might be achieved by exposing the person more to the English language. But at some point improvement by this means begins to level off, and further improvement will occur at best slowly, or not at all, if we do not manipulate the underlying capacity to speak a natural language which presumably is biologically based. Most of us would not acquire the mastery of English of a William Shakespeare or a James Joyce, however much we study the language. Nevertheless, biomedical means are not by themselves sufficient to give someone mastery of English, though they could be necessary for the attainment of certain levels of linguistic competence, or **(p.117)** for the attainment of these levels more quickly. This is the sort of role that we think that

biomedical means of enhancing altruism and the sense of justice could play if effective such means are discovered.

Education or instruction about what is morally good is not sufficient for moral enhancement because to be morally good involves not just knowing what is good, but also being so strongly motivated to do it that this overpowers selfish, nepotistic, xenophobic, etc., biases and impulses. An instructive comparison could be to people who know perfectly well that they ought not to smoke tobacco or eat sugary and fatty food because it is hazardous to their health, and yet do so out of weakness of will. Techniques of cognitive psychotherapy, such as vividly imagining how awful the harmful consequences of this behaviour could be of help, but might not suffice in all cases. Some individuals might be so strongly genetically disposed to nicotine or sugar addiction that they cannot get out of their addiction without biomedical treatment.

It might be wondered how such treatment could affect our attitudes without affecting our cognitions, but there are everyday experiences that could serve as models. We know that, by overexposure, we could get tired of a sensation or taste that we have hitherto liked and stop liking it, though the sensation or taste remains the same—the phrase ‘getting tired of a sensation or taste’ implies that the sensation or taste remains the same. Likewise, when people are cured of their irrational fear of spiders, by being exposed first to very small spiders, and successively to larger and larger spiders, until they are able to bear contact with spiders of a size much larger than the ones that used to occasion fear, without feeling fear any longer. There is no reason to think that their perception of those spiders have changed, though their emotional reaction to it has. It is conceivable that a pharmaceutical could cause the chemical changes in our brains underlying these attitudinal changes without our having to undergo any protracted exposure to the relevant stimuli, and suffer the stressful effects that such exposure might bring along. Pharmaceuticals are in fact given to people with agoraphobia or social phobia to reduce their anxiety.

The fact that being moral is not just a matter of possessing some knowledge is, as we have already tried to explain, the reason why the big chasm between our moral and technological capacity has opened up. Theoretical knowledge can be imparted from one generation to the next; thus, it will gradually accumulate over generations, making scientific and technological progress possible. But when people undergo great moral **(p.118)** development in the course of their lives, their moral competence will largely die with them. It cannot be transmitted to the next generation as easily as, say, mathematical competence. Nor could moral competence be used to construct devices that will help future generations to be moral in the way mathematics have been used to construct calculating devices. Moral competence is rather like artistic competence. Like artistic progress,

moral progress over the generations is considerably slower than technological progress.

Turning now to actual prospects for moral bioenhancement, one of the most promising lines of research has been on the hormone and neurotransmitter oxytocin. Oxytocin is naturally elevated by sex and touching, but it can also be elevated by nasal spray. It facilitates birth and breastfeeding in humans and other mammals, but it also appears to mediate maternal care, pair bonding, and other pro-social attitudes, like trust, sympathy and generosity (Insel et al., 2004).¹⁰ That is why it has been nick-named ‘the cuddle hormone’. When oxytocin is administered via nasal spray, it crosses into the brain. Several commonly used drugs are also thought to affect the release or metabolism of oxytocin. For example, the combined oral contraceptive pill, currently used by over 100 million women worldwide, is associated with elevated baseline oxytocin levels and is believed to increase oxytocin secretion. Similarly, glucocorticoids, widely used to treat asthma and other disorders of inflammation, are thought to modulate both the release of oxytocin and the expression of oxytocin receptors in some parts of the brain.

Kosfeld and collaborators investigated the relationship between oxytocin and trust in a simple game of cooperation (Kosfeld et al., 2005). Research subjects were divided into pairs and the first member of the pair (the ‘investor’) was asked to choose an amount of money to give to the second member (the ‘trustee’), knowing that the second member will receive three times the amount of money given. The second member then chooses an amount of money to return to the first member. The initial payment can thus be viewed as a signal of trust, while the return payment **(p.119)** can be interpreted as an indication of trustworthiness and gratitude. A greater level of trust signalled by the investor increases the total amount of money to be allocated between the two players, but the investor benefits from this only to the extent that the trustee is trustworthy and grateful. Prior to playing the game, participants were randomized to receive a nasal spray containing either oxytocin or placebo. Investors administered oxytocin exhibited significantly more trusting behaviour—that is, they entrusted the trustee with a significantly greater amount of money.

In a similar game to that to used by Kosfeld, Zak and associates found that the perception of a sign of trust by the trustee is accompanied by a spike in oxytocin levels and that the degree of trustworthiness exhibited by the trustee is positively and significantly correlated with the oxytocin level (Zak et al., 2004). Thus, in a population with universally elevated oxytocin levels increased trusting behaviour seems to be matched by increased trustworthiness.

However, oxytocin's effects on trust and other pro-social behaviour towards others appears to be sensitive to the group membership of these others. A research team led by de Dreu presented participants who had been randomized to receive either oxytocin or placebo via nasal spray with moral dilemma scenarios in which one individual would have to be sacrificed in order to save a greater number (de Dreu et al., 2011). Participants administered oxytocin were significantly more likely to sacrifice a different-race individual in order to save a group of race-unspecified others than they were to sacrifice a same-race individual in the same circumstances. In participants who had been administered a placebo, the likelihood of sacrificing an individual did not significantly depend on the racial group of the individual. The suggestion is that the pro-social effects of oxytocin may be limited to in-group members and exclude out-groups.

Further experiments by de Dreu's group indicated that oxytocin can also *reduce* pro-social behaviour towards out-group individuals where this helps one's in-group. Administration of oxytocin prior to participating in a group-based financial game induced 'tend and defend' reactions: it increased trust and cooperation within groups, but also increased non-cooperation with—though not offensive aggression against—members of other groups when this helped to protect one's in-group (de Dreu et al., 2010).

(p.120) This work supports the hypothesis that the pro-social effects of oxytocin are more accurately characterized as 'pro-in-group' effects, since the hormone can in fact induce antisocial behaviour when this conduces to the interests of one's in-group. Thus, it might be that a higher level of oxytocin amplifies the intensity of trust and reciprocity within an already favoured group rather than extends their range to out-groups. Since in-group favouritism seems to drive class and racial discrimination, which in extreme cases manifests itself in genocide and terrorism, administration of oxytocin would not by itself be an effective cure against these evils. It would have to go hand in hand with reasoning which undercuts race, sex etc. as grounds for moral differentiation. But that oxytocin by itself does not suffice for requisite moral enhancement does not show that it cannot be an indispensable aid.

To proceed to another class of current pharmaceuticals which have moral effects, selective serotonin reuptake inhibitors (SSRIs) are commonly prescribed for depression, anxiety, and obsessive compulsive disorder. They help govern activities such as eating and sleeping, and sexual activity. Millions of people world-wide use these drugs. SSRIs work by slowing the reabsorption of serotonin, a neurotransmitter crucially involved in mood, thereby making more of it available to stimulate receptors. Now SSRIs seem to make subjects more fair-minded and willing to cooperate. Tse and Bond (2002) had subjects play the dictator game—a game in which a dictator decides how a certain sum of money is to be divided between him or her and another participant—and found that

subjects administered the SSRI citalopram divided the sum more fairly than controls. Conversely, depletion of a precursor of serotonin, tryptophan, which would lead to reduced levels of serotonin, brought along lower rates of cooperation in the prisoners' dilemma game (Wood et al., 2006). The effect was only evident for subjects with depleted tryptophan in the first round of testing, suggesting that serotonin contributes to establishing a cooperative pattern of response, not maintaining it.

In the ultimatum game previously described, normal human subjects typically reject offers they regard as grossly unfair, despite the fact that rejection decreases their pay-off (in a one-shot game). Crockett and colleagues (2008) found that depletion of tryptophan led to increased rates of rejection of unfair offers relative to controls. This suggests that SSRIs may make subjects easier to exploit by modulating their assessment of what counts as (unacceptably) unfair. However, it is not clear how an **(p.121)** increased rate of rejection of unfair offers is to be interpreted: does it really signify a heightened sense of (un)fairness, or just greater aggressiveness and irascibility, which increase the probability that people will actually protest against what they see as unfairness? In any case it is clear that modifications of the brain by drugs like SSRIs have moral consequences.

The example of oxytocin and serotonin both show that manipulations of biology can have moral effects. There are then prospects of moral bioenhancement, even if so far no biomedical means of moral enhancement with sufficiently precise effects have been discovered, and perhaps they will never be. However, it is not surprising that no straightforward moral enhancers have hitherto been discovered because research into moral enhancement is a tiny field that is only a few years old.

Even if such means were discovered, the daunting task of applying them to a sufficient number of people—probably in the range of hundreds of millions—would remain. In any event, we are not envisaging that moral bioenhancement will ever reach a point at which traditional methods of moral education—or other social strategies like institutional redesign using incentives—will be redundant. As already explained, we think that these methods will need to be used as well and, indeed, that they should be employed more extensively than they are today. We are here highlighting biomedical means of moral enhancement because many people reject them for bad reasons, such as that they inevitably undercut freedom and responsibility. If effective, biomedical means could presumably be employed in a fashion that they undermine freedom and responsibility, but so could more traditional means if the application of them is so intense that it amounts to brainwashing.

In our view it is a serious mistake to reject moral bioenhancement out of hand because the need for human moral enhancement is so acute that we should not write off any potentially effective means without thorough examination. Significant moral enhancement of the human species appears to be necessary in order to ensure the survival of human civilization in the longer run. But many of us are loath to acknowledge that we are in need of moral improvement; it hurts our pride to acknowledge our moral deficiencies and, as consequence, to shoulder a possibly burdensome duty to rid ourselves of these deficiencies. It is more convenient to believe that the solution to the overwhelming problems of our time that we have outlined in this book is external to us and could be found either in technological inventions or political institutions like democracy. But **(p. 122)** these external instruments cannot handle the threats to the future of our civilization, unless they controlled by morally responsible people.

Although a great deal of moral improvement is hard going, it should be noted that some of it could be achieved quite easily because human beings are so prone to conformism. It is likely that merely by letting children grow up in a more altruistic environment they will become more altruistic. Indeed, it has been found that people are so readily influenced that just making them perform the task of unscrambling sentences about helpfulness increased their tendency to perform low-cost altruistic acts, such as picking up dropped objects (Macrae and Johnston, 1998). Certainly, such effects are only temporary. By contrast, having children grow up in a more altruist society, which discourages preferential treatment of friends and relatives to a greater extent than is present in liberal societies, is likely to have a lasting effect upon their altruistic proclivities. Thus, if a spirit of altruism began to spread in a community, this process could accelerate leading to societies with progressively more altruistic norms.

However, liberal democracies might be opposed to the implementation of more thorough-going programmes of moral education because it is at odds with the ideal that the state should strive for ideological and evaluative neutrality. Liberalism is based upon the doctrine of negative rights which, as we have suggested, we are designed by evolution to endorse. The main function of the state is taken to be to guard these rights. For instance, this doctrine of rights appears to be behind J. S. Mill's famous principle of liberty, according to which the only legitimate reason for state inference is harm to others, benefits to the self being insufficient to justify this measure. If injuring *any* interest of another were to count as harm, this principle could legitimize state interference against virtually all actions. So, Mill suggests a restriction to 'certain interests which, either by express legal provision or by tacit understanding, ought to be considered as rights' (1859/1978: 73). These interests will presumably turn out to be the ones that common sense takes to be protected by negative rights. According to the argument in Chapter 4, this would disqualify the contravention of all interests, which results in merely 'belief-mediated' distress.

But traditional liberalism has been too permissive as regards letting citizens of affluent societies adopt ways of living that waste the resources of the planet. We have suggested that so far all viable societies have inculcated something like the common-sense morality that we outlined in Chapter 2. This morality might have been conducive to the good of **(p.123)** these societies through preceding human history. However, it is too restricted in the current globalized setting of societies with advanced technology. This setting necessitates the inculcation of norms that are conducive to the good of the world community of which these societies are an integral part. Since such a revised morality goes beyond the morality to which we are naturally inclined, moral training will have to be more thoroughgoing and pursued intensively in school from the start.

It might be objected that we are proposing to imprint upon the public a morality that is philosophically controversial. Certainly, our scepticism of moral rights and causally-based responsibility is philosophically controversial. But it is not controversial to think that the limitation of our altruism to those who are near and dear, the bias towards the near future, and the numbness to larger numbers of sufferers are unjustifiable. The morality that we are proposing is to this extent a rather modest extension of common-sense morality, an extension which puts greater emphasis upon duties that common-sense morality already recognizes, in particular the duty to benefit those in need. The moral enhancement that we are recommending is largely a matter of motivating ourselves to do what we already believe to be right, of overcoming our moral weakness of will. However, if it is not accepted that that which we single out as morality is all of what remains of what is commonly called morality when it has been subjected to rational criticism, we could resort to stipulation and say that what we single out are those elements of morality—i.e. those doctrines which are about the well-being of other beings—that are of particular importance to the solution of the global problems that we have presented.

It should however be frankly admitted that moral bioenhancement worthy of the name is practically impossible at present and might remain so for so long that we will not master it, nor succeed in applying it on a sufficient scale, in time to help us to deal with the catastrophic problems that we have outlined. But our point is just that the predicament of humankind is so serious that all possible ways out of it should be explored. Therefore, it is important that moral bioenhancement is not written off without good reason. Because of the gravity of the current human predicament, effective moral bioenhancement, were it technically feasible, would in our view be the most important kind of biomedical enhancement.

However, it must not be forgotten that the techniques of moral bioenhancement raise the same moral problems that all powerful technological innovations create: how to ensure a wise and proper application of them. **(p.124)** All technology is liable to the dual use problem: they can be put to both beneficial

and harmful uses. In the case of techniques of moral bioenhancement, this takes the form of a bootstrapping problem: it is human beings, who themselves need to be morally enhanced, who have to (a) be enough interested in being morally enhanced to set aside sufficient resources for research into biomedical means of moral enhancement, and (b) if effective biomedical means are discovered, to make a morally wise use of them. We see no reason to think that this research need be so costly that (a) would be a problem. In our opinion it is (b) which presents the greatest problem: is to hope for a wise use of biomedical means of moral enhancement not to hope for too much when humans have made such unwise use of so much scientific technology? We have already warned against the tendency to be overconfident about what we can achieve in the future by yet to be discovered means; so, we must be careful not to pin our hopes too high. But, on the other hand, there is the opposite risk that too much pessimism about the possibility of moral bioenhancement could lead to it being prematurely dropped from the agenda.

Morally enhancing the majority of people in modern democracies is certainly a huge task. But we are not assuming that such an enhancement will have to be accomplished by biomedical means alone. Traditional moral education also has a part to play, perhaps the largest part. These means of moral enhancement could interact with the impact of social and political reforms that have greater chances of gaining democratic voters' support once there has been some measure of moral enhancement. Thus, we are imagining an interplay between biomedical and social/political techniques rather than the former being alone in the driver's seat. But, since the discovery of effective techniques of moral bioenhancement still lies far ahead, it is difficult to envisage what form a large-scale application of them should assume. And, indeed, it might turn out that such an application is never needed because it has been pre-empted by other means; traditional means of moral enhancement or institutional means (or because human civilization has been ruined). With respect to some of the obstacles to moral behaviour that we have discussed, like the conception of responsibility as being causally-based, the bias towards the near and the availability bias, we are not familiar with any other kind of remedy than a traditional, cognitive one.

Irrespective of whether the means of moral enhancement be traditional or biomedical, they will scarcely be enough, given the gravity of the **(p.125)** problems that we have described: it is likely that we will also have to accept a rather extensive surveillance by the state, since there will inevitably be a few individuals among us who are bent upon using the powerful means of scientific technology to wreak havoc. Moral enhancement could not realistically prevent there being a small number of morally warped individuals deploying powerful technology for nefarious purposes. Therefore, we think that the application of means of surveillance of citizens which go beyond those used in the fight against traditional crimes are necessary. These means involve setting aside what people in liberal democracies have come to regard as rights, in particular the right to

privacy. Likewise, the freedom of media needs to be restricted with respect to the publication of scientific studies that could supply terrorists with horrific weapons, as illustrated by the mousepox case discussed in Chapter 4. But moral education also has a role to play in this context, for instance, to combat xenophobia which might flare up against ethnic groups some of whose members have committed acts of terrorism. However, some shortcomings of a cognitive sort need to be corrected, in particular the availability bias. Otherwise, people will be inclined to exaggerate the risk of future terrorist acts of a kind that they have already experienced and be blind to the possibility of other kinds.

It seems nowadays to be a common assumption that science could provide a cure for more or less all of the serious problems that humanity faces. This belief is encouraged by the fact that, thanks to science, in affluent societies more people lead longer and better lives than ever before in human history. But there is a non-negligible risk that, as science probes deeper and deeper into nature, science will unleash some highly destructive, uncontrollable processes. Rees (2003) mentions the risks of powerful particle accelerators, like the one in the CERN laboratory in Geneva, of runaway gene-modified organisms and self-replicating, omnivorous nanomachines. Nanotechnology illustrates the dilemma well. On the one hand, there is the hope, for instance, that it could play a central role in counteracting global warming by creating nanomachines which devour carbon dioxide molecules in the atmosphere. On the other hand, this could let loose uncontrollable, self-replicating, and omnivorous nanomachines.

So, scientific progress undoubtedly generates catastrophic risks. But if it were brought to a halt—this certainly stretches the imagination—this might mean that we divest ourselves of the possibility of dealing with **(p.126)** other catastrophic risks, which already exist and which we could otherwise have disarmed. An example of such a kind of risk is the risk of asteroid strikes. Some 65 millions ago an asteroid hit Earth, and it is commonly believed to have rubbed out the dinosaurs and many other life forms. This asteroid is estimated to have been 10 km in diameter, and asteroids of this size or more are expected to hit the Earth only once in 50 to 100 million years. But collisions with smaller asteroids, with a diameter of 1.5 to 2 km, might kill a billion people or more, and they are expected to occur roughly twice in a million years (Posner, 2004: 25-6).

Approximately 75,000 years ago a volcano erupted in Toba, Indonesia. As a result of the tremendous quantities of sunlight-blocking ash it spewed out into the atmosphere, the global temperature dropped by 5–15°C. It is surmised that the human population shrunk to around 4,000 reproductive individuals (Rampino, 2008: 211-12). According to some estimates, such super-eruptions occur once in 50,000 years. There is therefore reason to try to predict them and to take precautions against them.

However, it seems clear that the prospect of reducing the risk of such mega-threats could not outweigh the mega-threats that scientific progress has generated. In terms of mega-threats the balance comes out against contemporary science and technology. Rees judges that ‘the odds are no better than fifty-fifty that our present civilisation on the Earth will survive to the end of the present century’ (2003: 8). Such an estimate would have been wildly implausible with respect to any other hundred year period before 1950s, the time at which humans acquired the nuclear capacity to blow up the Earth.¹¹ At that time, the power to cause Ultimate Harm lay in the hands of only a few; today it is in the hands of many more, and the number is likely to grow. Consequently, it seems indisputable that contemporary scientific technology has increased the risk of world-wide catastrophe, even if it is the case that Rees’s estimate of the risk is somewhat exaggerated. Worse still, if the progress of scientific technology continues, and there is no moral enhancement of human beings, the probability that civilization will survive not just the present century, but also the following centuries will be progressively lower. This is a horrifying trend that must be broken.

(p.127) We are inclined to believe that at the time, half a century ago or so, when scientific technology provided us with means of causing Ultimate Harm, technological development reached a stage at which it became worse *all things considered* for us to have the current means of scientific technology, given that we are not capable of handling them in a morally responsible way. If life on this planet were to end soon in a catastrophe caused by modern technology, when it would otherwise have continued for millennia, the final judgement will have to be that the present technological development has been for the worse all things considered. It is possible that there was a turning point in the development of scientific technology, that at some point—of course not a very precise point—it turned from being for the better all things considered to being for the worse all things considered. We believe that such a turning point could have been when it provided us with means of causing Ultimate Harm. Thus, we do not wish to commit ourselves to the extreme claim that technological development *from its inception*, in Stone Age, say, has been for the worse all things considered.

Nonetheless, this extreme claim might not be so grotesquely implausible as it might seem at first sight. For instance, Craig Dilworth writes:

it may be suggested that the Upper Palaeolithic (40,000–25,000 BP) constitutes the high point in the human way of life to date...it can fairly be said that we never had it so good before, and we’ve never had it so good since. Though average longevity was short by modern Western standards, those who survived infanticide and death related to protowar lived to an advanced age, 60 to 70 being quite possible. (2010: 204)

Dilworth notes that people at this time, at the beginning of human technology, were well nourished, since big game was plentiful, and suffered very few diseases. Most of our infectious diseases come from domesticated animals; according to Dilworth, ‘humans now share 65 diseases with dogs, 50 with cattle, 46 with sheep and goats, and 42 with pigs’ (2010: 244). At the time in question the domestication of animals had not yet started. Nor were these hunter-gatherers afflicted by diseases related to tobacco smoking, obesity, and pollution, which claim many lives today.

Consequently, if we compare the *average* sum of welfare of human life, it may not be much higher today, for although many people in contemporary affluent societies have a higher quality of life than Cro-Magnons—because they enjoy the blessings of modern technology **(p.128)** and culture—at least as many in developing countries lead lives that are of lower quality because they starve and are plagued by a plethora of diseases. What is striking, then, is that the extraordinary technological advance has done comparatively little to raise this average. There are at least two reasons for this. First, the profusion of material wealth that this advance has generated has created large problems of distribution, which we have been incapable of handling. Secondly, a considerable portion of the products that this advance has yielded serves the function of remedying defects that this advance itself has given rise to. This is consonant with our claim about the urgent need for moral enhancement, in order to ensure a morally wiser use of this advanced technology.

Of course, if we consider the *total sum* of human welfare instead of its average level, it is vastly greater today than in the days of Cro-Magnon, since the human population is now more than a thousand times bigger. But the huge amount of current human welfare might be bought at the price of less human welfare in the future, since we are now depleting the resources of the planet.¹² And in any case, the goodness of an outcome is determined not just by the total sum of its welfare—as the so-called repugnant conclusion brings out (see Parfit, 1984: ch. 17)—we need also to consider its quality, as well as the justice of its distribution. The general point is, however, just that it should not be assumed that average improvements as regards human life quality or welfare automatically march in step with technological progress. It should not be taken for granted that just because the technology of a society is primitive the average quality of life in it is poor.

However, in one respect there has been clear change to the better in the course of human history: the percentage of people who suffer violent deaths—and from other acts of violence such as torture and rape—has gone down considerably since prehistoric times, with the switch from non-state to state societies and the subsequent growth of commerce and trade.¹³ Hence, Dilworth’s caveat ‘those who survived infanticide and **(p.129)** death related to protowar’.¹⁴ The decline of the proportional number of acts of violence is a necessary accompaniment of

the enlargement of human societies from around 100 members to millions and even billions of members. Such huge societies could not exist without an authority in possession of a monopoly on violence, which effectively curbs the aggressiveness of their citizens. Furthermore, the thick web of trade and commerce inside these societies and between human societies all over the world has produced a level of welfare that people are reluctant to put at risk by starting violent conflicts. International commerce, alongside the devastating power of modern weapons—both fruits of advanced scientific technology—act as powerful war deterrents, by boosting the costs of waging war.

But we cannot rest assured that this pacification process will continue. A threatening depletion of natural resources could raise the prospect of gain by predatory wars sufficiently to make it worthwhile in the eyes of national leaders. Moreover, religion continues to pose a risk because by promising rewards in an afterlife it could justify wars which in terms of this world are ruinous for all parties. Perhaps it is less likely that nation states will wage wars on purely religious grounds, but then we should keep in mind that, if not already in the present, at least in the near future, small groups or even single individuals may be in possession of devastating weapons of mass destruction, and they might well be fanatical enough to put them to full use, though from a worldly perspective they, along with everyone else, will lose out.

Thus, against the decreasing incidence of violent outbreaks, on individual and national levels, we must set the enormously greater destructive potential of an individual instance. The number of humans who die at the hands of their fellows is a function not only of (a) how many times their fellows engage in acts of killing, but also of (b) the effectiveness of the weapons or means of killing that they then use. We agree that, proportionally, the number of occasions when humans engage in acts of killing has gone down, but the weapons that they have at their disposal are more effective than ever and are likely to be even more effective in the future. So, we cannot be confident that, because people in general have become less prone to indulge in killings and other forms of violence, **(p.130)** the percentage of victims of human killing will remain comparatively low as e.g. the victims of rape—where the (b)-factor is negligible—presumably will.

However, irrespective of what the upshot of this risk assessment of the development of science and technology is, it is beyond question that this development has made our moral responsibility, including our responsibility for handling the risks that we encounter, larger than it has ever been. An age of powerful scientific technology is inescapably also an age of wide-ranging moral responsibility; with power of action comes responsibility. Once we gain the power to alter nature's course, we become responsible for allowing nature to take the course it takes. Thus, there is a moral gulf between accepting what nature delivered when there was nothing that we could do about it and

accepting what nature delivers when we can affect the course of nature: in the latter case we are responsible for letting it happen and must morally justify our stand.

The fact that it is easier to harm than to benefit accentuates how important it is that our behaviour is under moral control, as well factually informed. As our powers of action are enlarged by scientific technology, our capacity to harm grows more than our capacity to do good. In the pre-scientific past many of the risks with which the world presented us were ones that we could not do anything about. Now we can to a considerable extent do something about the risks we face, by applying knowledge that science yields. Thus, we are able to do a lot of good, but we might be responsible for even more harm, by active designs or omissions, since it is easier to harm than to benefit. In a democracy the responsibility to decide ultimately falls upon the shoulders of the voters, and they should be morally fostered and scientifically informed to carry it well. Wise decisions require not only good scientific knowledge but also internalization of a robust and well-grounded set of moral values.

There is an explosion of possibilities of scientific research, and it is getting increasingly important to be keenly selective and give priority to research which is most beneficial to human beings and other organisms on Earth. As things stand at present, research is rather directed by the interest of the most privileged to have further economic growth and greater affluence. This is leading to a depletion of natural resources, a reckless release of waste products, and a widening of the gap between the best and the worst off people. What is needed is what might be called a **(p.131)** 'science-sophy', moral wisdom as regards the pursuit of scientific research and its practical applications. Alongside moral judiciousness, this wisdom involves a good measure of scientific knowledge: in a modern democratic society in which political decisions involve a lot of science, it is desirable that the general public and politicians possess a reasonable knowledge of science. According to such a science-sophy, the clearest example of scientific research that we should not engage is probably military and armament research, which consumes gigantic resources, even in the poorest nations, such as North Korea.

A different sort of example of arguably misguided research might be research into the possibility of extending human longevity beyond the 120 years or so now thought feasible. Such research tends to increase the acute population problems that already exist, and to enlarge the huge gap in life-expectancy that already obtains between people in the developed and the developing nations. A science-sophy would recommend moral bioenhancement rather than any other kind of enhancement. Generally speaking, scientific research should be informed by a global, number-sensitive altruism and sense of justice which is not temporally biased, or reigned in by a conception of responsibility as causally-based. By contrast, much contemporary research remains governed by the selfish interests

of the rich, which aggravates global inequality and harms the interests of future generations.

It is a naïve illusion to think that we could eventually rid ourselves of the necessity of having to make morally hard decisions with respect to science because it will in the future enable us to do everything we want. Even if, as is most unlikely, science were to develop, for example, to a point at which it would allow us to take care of all present climatic and environmental problems (at least in so far as they affect human welfare), without our having to restrain our consumerist lifestyle, or radically reduce our number, there would still be the moral problem of how to handle the risks of intentional misuses of this science which will have to be exceedingly powerful. In general, scientific progress could not relieve us of moral responsibility; instead, it inevitably extends its range. Any powerful technology is liable to the dual use problem, and since it is easier to harm than to benefit, it is most likely that it can be misused to create greater harm than it can do good.

Consider, for instance, a geoengineering scheme aimed at keeping the temperature of the Earth down by increasing the extent to which its **(p.132)** atmosphere reflects incoming solar radiation. This could be done by injecting sulphur dioxide gas into the stratosphere to create sulphate aerosols, particles that reflect solar radiation. Carbon dioxide could then accumulate in the atmosphere without any temperature rise. While this scheme would counteract a temperature increase, it would have an opposite effect on the acidification of the oceans which is due to their absorption of carbon dioxide from the atmosphere. This increasing acidification would lead to a dissolution of coral reefs. According to some conjectures, the injection of sulphur dioxide into the stratosphere would also disrupt the Asian and African summer monsoons, which could jeopardize the food production for billions (Hamilton, 2010: 177). However, the most worrying aspect of such an ambitious piece of climate engineering is that, as a large-scale measure that has never been tried out, it is bound to have unknown side effects, and these are more likely to be harmful than beneficial, since it is easier to damage a functioning system than to improve upon it. But once we have started to inject sulphur dioxide in the stratosphere to keep the temperature down, we are trapped: if we observe untoward effects, we cannot discontinue these injections without causing a devastating temperature jump (Hamilton, 2010: 182). Moreover, we have not got around the difficulty of securing international agreements. Different nations are likely to want to set the global 'thermostat' differently, e.g. China is likely to want to set it lower than Russia. Consequently, serious international conflicts over the 'ideal' global temperature might result (Hamilton, 2010: 182-3).

All the same, since so little is currently done to reduce the emission of greenhouse gases, we must calculate with a situation in which humanity will resort to large-scale geoengineering to avoid climate disaster. It is desirable that

we are well prepared for such an emergency by having as carefully as possible investigated the pros and cons of different techniques of geoengineering. Some have objected that if we open up the prospect of geoengineering, this encourages the current slackness about reducing the emission of greenhouse gases. In our opinion, however, this slackness is already so great that the risk of increasing it somewhat cannot do much damage.

However, we believe that moral enhancement, by traditional means as well as novel means that biomedical research is likely to unearth, has priority. This is not just with respect to bringing about a willingness to stop deleterious climate change. We have seen that we need a very **(p.133)** advanced technology—in all probability a more advanced technology than we already possess—in order to provide the enormous, and increasing, human population with a decent standard of life without exhausting the resources of the planet. But it is vain to hope for a technological fix that by itself solves this equation. Without moral restraint it is likely that, as has happened in the past, a more efficient technology will be spent on a further expansion of human activities (a tendency elaborated at great length by Dilworth, 2010). Moreover, since it is easier to harm than to benefit, more efficient technology will bring in its wake a greater risk of Ultimate Harm. Thus, we face a dilemma with respect technological progress which only moral enhancement can take us out of: we need it to improve the lot of humanity, but it brings along a risk of Ultimate Harm. In our view, moral enhancement is necessary if human civilization is to have a reasonable chance of surviving not merely the present century but also following centuries.

We shall not attempt to predict whether, by one means or another, liberal democracies will ever come to possess sufficient moral wisdom in the employment of scientific technology, or whether they will rather founder on the problems generated by this technology. Our main point is merely that liberal democracies are in need of moral enhancement in order to deal safely with the overwhelming power of modern technology. It is crucial that we be aware of the moral limitations of our nature, and do whatever we can to correct these limitations, by traditional or new scientific means. We are not trying to predict to what extent we shall in fact succeed in rectifying these limitations in time, or what the future of humanity will in fact be. This is because we are of the opinion that the future of humanity cannot be reliably predicted; at best, we can predict roughly what is likely to happen *if* various policies are adopted. As already remarked, this is because we have the capacity to overturn categorical predictions that we make by making decisions on the basis of them. But even conditional predictions of larger scale outcomes are highly unreliable because these outcomes depend upon innumerable small factors, each of which could have big effects on the future. Imagine, for instance, how different the world might have been if Al Gore instead of George W. Bush had been declared the winner of the tight US Presidential election of 2000.

In conclusion, more efficient technology seems necessary to provide the huge human population on Earth with an acceptable living standard without wearing out the planet. Nevertheless, since it is easier to harm **(p.134)** than to benefit, there is likely to be a turning point at which the growth of human powers of action by means of scientific technology becomes for the worse, all things considered, because the moral shortcomings of humankind make the risks of catastrophic misuses of these powers too great. These risks are real because human psychology and morality are adapted to life in small, close-knit communities with simple technology, not the societies with millions of citizens and an advanced scientific technology that we find today. Simply because the human population is larger than ever, immoral actions occur more frequently today than ever, and with potentially more disastrous consequences because of the enormous number of agents and the means of modern technology.

We believe that a turning point was passed at least some fifty years ago when humans acquired the means of causing Ultimate Harm by nuclear weapons. However, it is possible for humankind to improve morally to the extent that the possession of the overwhelming powers of action supplied by scientific technology could be used to create an unprecedented amount of human—and animal—welfare. This is a definite possibility, not least because biomedical techniques that could be provided by the advance of science, along with the techniques of traditional moral education, could be employed to promote a moral enhancement of humankind. But we are not in the business of trying to predict whether or not a happy ending is likelier than a catastrophic outcome, nor what, if anything, will in the end be the most effective means to such a happy ending.

Notes:

⁽¹⁾ For further discussion of the unpredictability of decisions, see Persson, 2005: ch. 31.

⁽²⁾ There is an especially great risk that the interests of non-human animals, existing in the present or future, be set aside since, apart from the fact that they have no voting power, the moral weight of their interests is more contested. The moral status of non-human animals is an issue that we cannot discuss here, but this should not be taken to imply that we regard it as unimportant. However, since the moral status of non-human animals is contested, we do not want the case for sustainable policies upon it. For a discussion of the threat to biodiversity, see e.g. Wilson, 2002.

⁽³⁾ The Nazi ideology is too familiar to need mentioning; it is more noteworthy that an Oxford philosopher, Hastings Rashdall, a little over a hundred years ago could write something as blatantly racist as this: ‘the lower well being—it may be the very existence—of countless Chinamen or negroes must be sacrificed that

a higher life may be possible for a much smaller number of white men' (1907: 238-9).

(⁴) For another defence of moral bioenhancement, see Douglas, 2008. For a defence of biomedical enhancement in general, see e.g. Buchanan, 2010.

(⁵) Contrast what for instance Jonathan Haidt, 2003, calls 'moral emotions'. These include for instance 'other-condemning' emotions like anger, contempt, and disgust. Obviously, these can by themselves issue in behaviour against the targeted individuals that is morally wrong. cf. also what Jesse Prinz and Shaun Nichols take to be moral emotions (2010). But even if we do not regard, e.g. anger, as a specifically moral emotion, we do not deny, of course, that a reduction of anger can be a moral improvement.

(⁶) Cf. what de Waal calls 'preconcern': 'Preconcern is an attraction toward anyone whose agony affects you. It doesn't require imagining yourself in the other situation' (2010: 96). As we do, de Waal takes genuine sympathetic concern to encapsulate empathy (2010: 88).

(⁷) But only roughly, since Baron-Cohen thinks that sympathy is only one example of the affective responses that empathy encompasses (2003: 26-7). This is a disagreement that we need not sort out here.

(⁸) Cf. de Waal: 'I'd be reluctant to radically change the human condition. But if I could change one thing, it would be to expand the range of fellow feeling. The greatest problem today, with so many groups rubbing shoulders on a crowded planet, is excessive loyalty to one's own nation, group, or religion' (2010: 203).

(⁹) For a fuller reply to Harris, see Persson and Savulescu, 2011. See also the reply to Harris by Douglas, forthcoming.

(¹⁰) Here we cannot resist quoting the sagacious remarked by Tsutomu Yamaguchi who survived the atomic blasts both at Hiroshima and Nagasaki: 'The only people who should be allowed to govern countries with nuclear weapons are mothers, those who are still breast-feeding their babies' (D. Garner: 'After the atom bomb's shock, the real horrors began unfolding', *New York Times*, 20 January 2010.)

(¹¹) It should also be remarked, that around 1950, the human population on Earth was still only 2.5 billion, so the catastrophic degradation of the environment had not yet gained momentum.

(¹²) In this respect there is a parallel between the present time and the period Dilworth describes because the welfare of the Cro-Magnons came at the expense of the mayhem they inflicted upon the mega-fauna of newly colonized territory. Although, as Pinker notes (2011: 454-74), there has recently been progress in the recognition of animal rights or welfare; the expansion of the meat industry

and the loss of biodiversity mean that humans kill animals at a faster rate than ever.

(¹³) As forcefully argued by Pinker, 2011: esp. chs. 2 and 3.

(¹⁴) We bracket the reduction of infanticide, since it bears as little on the issue of pacification as does the high percentage of pregnancies that nowadays end in abortion.

Access brought to you by: