

Kent Academic Repository

Full text document (pdf)

Citation for published version

Robinson, Alexandra Jane (2021) Ethical Transhumanism: How can a nudge approach to public health make human enhancement more ethical? Doctor of Philosophy (PhD) thesis, University of Kent, University of Kent.

DOI

Link to record in KAR

<https://kar.kent.ac.uk/92731/>

Document Version

UNSPECIFIED

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:

researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

**Ethical Transhumanism: How can a nudge approach to public health make human
enhancement more ethical?**

Alexandra Jane Robinson

Supervised by Professor Simon Kirchin and Dr Lauren Ware

Words: 70,109

Department of Philosophy, University of Kent

Submitted on the 19th of September 2021 for the Degree of Doctor of Philosophy

I am extremely grateful to my supervisors, Professor Simon Kirchin and Dr Lauren Ware, for their advice, support, and endless patience.

With thanks to Professor Lisa Bortolotti and Dr David Corfield for their consideration and invaluable feedback as Examiners, and to Dr Lubomira Radoilska for her support and encouragement of earlier research which made this thesis possible.

I am indebted to Karolina for her help in proofreading, and to Mausel the cat without whom many productive early mornings might have passed me by.

For my parents, who taught me to keep going.

Abstract

Transhumanism at once embodies our most modern thinking and our biggest longstanding problems.

Transhumanism aims to enhance human core capacities: health-span, lifespan, and cognition. The thesis answers the following ethical challenges arising from transhumanist aims. First, whether transhumanism can be an ethical endeavour if it relies on authoritarian intervention by governments and governing bodies to change, generate and enforce behaviour, or to influence and enforce the uptake of medical procedures. Second, the thesis answers the challenge that it is unethical deliberately to encourage the uptake of and pursuit of medical transhumanism given the extent of accessibility and distributive issues that remain unresolved in existing medicine. Finally, the thesis addresses a particular mental health crisis that is often predicted for transhuman beings, namely loss of meaning from loss of death and vulnerability, resulting in widespread loss of social cohesion. The thesis argues that the right solution to the first two problems is a libertarian paternalist approach, viz. nudging, and that this approach will also neutralise the risk of widespread and inevitable boredom or alienation that might otherwise result from the widespread introduction of human enhancement if people are nudged to engage more and more reflectively in their enhancement choices. Additionally, lifestyle issues like obesity, heart disease, cancers, and inaccessibility of vaccines and birth control pose unresolved problems for existing general medicine, killing millions every year worldwide. As a result, another serious challenge for enhancement medicine, which I propose would be addressed by the nudge approach, is to justify its place in the professional domain of medicine.

CONTENTS

Section One

1 Thesis Introduction	5
2 Human Enhancement and Health: Transhumanist Primer	19
3 Three Key Philosophical Dialogues	
First Dialogue: Autonomy and Behavioural Insights into Free action	35
Second Dialogue: How Medical Knowledge Changes Medical Practice	58
Third Dialogue: Emotions and Medical Advice	71

Section Two

4 Is Human Enhancement a Public Health Problem?	88
5 Organ Donation as a Key Example of Behavioural Intervention	121
6 Justifying Emotional Nudges in Public Health	150
Thesis Conclusion	169
Postscript about Covid-19 and the time of writing	174
Notes	176
Bibliography	191
Appendix	205

SECTION ONE

1

Thesis Introduction

Why does transhumanism need nudges? Why should governing bodies be interested in nudging towards transhumanism?

The first premise of the thesis is that a libertarian paternalist approach is well suited to the challenges faced by public health problems, where individuals need to act a certain way for some project to succeed or milestone to be met. Ethical challenges for public health include the need to promote and maintain the autonomy of individual agents whilst allowing the enforcement of public health policy across demographics within a society. Political challenges for public health include the need to maintain authority and trust in the oversight of the healthcare provision, whether it is nationalised or privatised, or a combination of available providers. Social challenges for public health include the need to motivate groups and to make sure the needs of diverse groups and diverse individuals are met, without privileging or dividing society on healthcare lines. Public health efforts should not be limited to only certain individuals, and the benefit of public health projects should not be only for the minority, but must prioritise the needs and triumph the values of most.¹ A well-established literature on behaviour and decision-making, more recently including psychological studies and sociological reports exploring many modern examples of public health measures in practice, shows the potential of a libertarian paternalist (also called 'nudge'² or 'behavioural insights'³) approach to public health policy. Nudges in public health have been used to improve outcomes in campaigns encouraging public uptake of medical interventions, like voluntary organ donation,⁴ 'don't drink and drive' campaigns,⁵ and infant vaccine programmes,⁶ in Europe and North America.

Regarding the increase of donor organs for transplant, the NHS Blood and Transplant strategy for the next decade⁷ draw on the work of Randhawa⁸ into racial disparities of accessibility and barriers to healthcare for racial groups in the UK. The strategy highlights the need to continue to address behavioural barriers to accessibility of services:

“Review of published evidence and engagement with the public and patients has shown that access to transplantation is not equitable for all patients. People from a BAME background have a high need for organ transplantation, while donation rates are generally lower. Barriers to organ donation in BAME communities restrict the availability of well-matched organs and results in relatively long waiting times for transplantation, with an increased risk of death on the transplant list. Understanding the barriers that lead to donor opt-out registrations and family consent among people from a BAME background is essential to match demand for transplantation and availability of suitable donor organs.”⁹

In the literature on transhumanism, challenges for ethical human enhancement medicine parallel the public health challenges for the kinds of public health policies that benefit from nudges. Ethical transhumanism must preserve autonomy whilst encouraging cooperative healthy behaviour. Transhumanists who aim to bring human enhancement into general medicine face a problem of ethical distribution and ethical enforcement, because upholding and maintaining a global standard of enhanced health would also rely on governing bodies striking the right balance of liberty and authority. Governments and governing bodies who seek to encourage and educate about the potential for human enhancement to change and improve the quality of human life must overcome uncertainty and fear of bodily interventions, and of equal distribution (fairness, justice, resentment).

I argue that nudges would benefit transhumanist activism in a significant way. Nudges could hasten the introduction of human enhancement technology into general medicine and public health in a way that would not undermine or detract from existing global health efforts to proceed carefully, in light of the above concerns for transhumanism as an idea. In the thesis, I engage with behavioural insights theory, nudges, which have been influential in recent policymaking and popular with governments in both the United States and the United Kingdom in recent years¹⁰ as a possible solution to concerns about ethical transhumanism. Some public health projects which seem similar to key examples of possible medical human enhancement, have achieved significant improvements in public health with the help of nudges, and there are comparable social, political and philosophical challenges.

Second, in cases where emerging technology stands to improve the general quality of life, positive intervention by governments or governing bodies at the right time could significantly improve novel application of technology and overcome the effects of the status quo bias, to reduce techno-stress.¹¹ Having public feeling on side should be an important measure of the success of an emerging technology or for emerging applications of technology.¹² This is

because positive public feeling significantly increases the potential for emergent technology to increase wellbeing, which should be as significant as the economic or straightforwardly innovative success of technology. Further, the literature on behavioural insights promotes the idea that success in the public eye, and the contribution of public agreement to wellbeing, should not be the subject of study only in hindsight.

Taking the same approach based on nudges, I argue it is possible and important to intervene early and carefully to influence public opinion, without undermining public consent for widespread (even global) public health policy changes. I suggest that the positive effects of technological change are not stifled by technostress or a tendency for people to prefer the status quo and to resist change, which strengthen autonomy and improve the possibility that individuals can participate in an informed way, in public health.

Third, global interventions to create policy and influence public opinion about enhancement medicine would help the progress and effect of transhumanism, but also would help improve and hasten a more general shift toward a more egalitarian and responsible global health policy for human beings. Critics of transhumanism suggest that disillusionment and disengagement with healthcare could easily result from human enhancement: alienation and a sharper divide between perceived 'haves and have nots'; individuals enhanced without their consent would be more likely to feel uneasy about their modified core capacities; unwanted long life could result in boredom; and enhanced individuals might stagnate, with established norms no longer fit for purpose to guide lifestyles.

In short, it is the absence of engagement and early social interventions that is most likely to result in the ethical challenges predicted of human enhancement medicine, not human enhancement medicine itself. When policy aims not only to follow but engage public cooperation and understanding of emerging technology, public disillusionment and disengagement with medical research and development seem to be less likely.

Although transhumanist goals like human immortality or habitation of other planets in the solar system might seem fanciful and might once have been more appropriate subjects for science fiction, there is increasing reason to believe they are sensible, practical answers to the question 'what can human beings do next?'. There is increasing clinical evidence that life extension and cellular damage repair is a realistic medical goal, which could lead to greatly elongated healthy lifespans for human beings. Transhumanists tend to suppose that immortality, although it may never mean total and invulnerable freedom from death, which does not seem desirable in any case, will inevitably become a reality as indefinite longevity.

The technological possibility of indefinite longevity, through cellular repair and gene editing, means it is more important now than at previous times that philosophy should address philosophical and bioethical concerns about how and when to introduce human enhancement into general medicine. Life extension, modification of intelligence, improved skeletal and muscular strength, resistance against deterioration and previously unmanageable genetic illnesses, direct brain interfacing with machines, and re-growth of lost body parts are more realistic now than they have ever been, but this does not mean they should automatically become everyday medicine as soon as they become technically feasible in the laboratory. Serious ethical concerns about the imminent matter of introducing enhancement into conventional medicine are being voiced. In medical journals, in reports by large healthcare organisations (for example, UK-based Nuffield Health¹³ and the Wellcome Trust¹⁴), in global public health advice and policy implementation organisations (such as the World Health Organisation), and from government agencies like the British Parliamentary Office of Science and Technology, in collaboration with The Hague¹⁵), the possibility of human enhancement amounting to transhumanism has been taken seriously since the 1970s.

In the last decade, the urgency and seriousness of the need for global consensus on ethical regulation of human enhancement has grown considerably, particularly following the announcement of successful human enhancement in China, near the end of 2018.¹⁶ Transhumanist organisations like SENS (Strategies for Engineered Negligible Senescence) and Óisín Biotechnologies, in the United States, have flourished in the last decade, and available funding for human enhancement and negligible senescence research was recently reported to have outstripped the uptake from start-up organisations in the industry.¹⁷ In light of this apparent boom-time for transhumanist research and human enhancement medicine, it has been suggested¹⁸ that questions such as, ‘what makes human beings uniquely human?’, ‘what is the limit of our human potential?’, and ‘should we undertake self-transformation or self-evolution, as a species?’, traditionally the preserve of philosophy and existentialism, and of clear interest to transhumanism, are coming more and more into the public sphere. These longstanding philosophical questions have been matters of vague public awareness for quite some time, particularly as more personal questions about self and determination, and particularly when popularised in the Existentialist philosophy of the last century.¹⁹ However, it does seem transhumanist, existentialist, philosophical questions about human being have re-entered the public consciousness as scientific and medical questions with new potential.

A significant part of the scientific community now seems to agree that CRISPR technology is promising as a means of improving health outcomes in the face of illness.²⁰ CRISPR has been beneficial in improving the prognosis for genetic issues in other species.²¹ In clinical research, including animal trials and early trials on human embryos *not* for implantation and pregnancy, over several years, CRISPR has appeared promising to the expert community in general, though predicting the future uses of the technology was difficult and sometimes controversial.²² Once it was reported, late in 2018, that CRISPR-Cas9 had been instrumental in successfully editing living human infants, however, the possibility of enhancement inevitably captured the public imagination anew. Just as it gave thought-leaders increased cause for concern and gave the scientific community reason to scrutinise the speed of endeavours to introduce CRISPR into general medicine, the public became newly aware of how close the advent of transhumanist human enhancement medicine might be.²³

But why should the public, governments, and global health bodies pay attention to transhumanism, as much as to cures for cancer or heart disease?

The Shenzhen human gene modification trial in 2018 was notable, not least because of its novelty and illegality. Philosophical questions were revived for the public as newly scientific, rather than as subjects for academic thought experiments, and commentary tended to ask whether it was *right* to enhance human beings using the technology. However close the technology appears to have come to allowing successful human genetic enhancement, doubts remain about the shortcomings of structures of medical authority, and whether professionals have the right to carry out enhancement. These seem to challenge both the aims of transhumanism and whether transhumanism can be achieved in acceptable, ethical ways. The availability of CRISPR-Cas9 technology raises questions about the limits of existing medical frameworks, which would need to change to account for the use of enhancement technologies, not least because disease-based frameworks for medical care do not allow for enhancement as medicine.²⁴ Doubts remain over the ethical implications, should any national governments seek to acquire and distribute enhancement technology into existing national healthcare systems to enhance their citizens where other nations decline to do so.²⁵

Existing defences of transhumanist ideals like life-extension and immortality, in the philosophical literature have tended to address concerns about the risk of disparity between 'haves and have-nots', alienation from *feeling human*, and about species divergence that might pose an existential threat to humanity. Bioethicists have asked about the autonomy of the enhanced,²⁶ and about the risks posed by boredom and loss of meaning to immortal

or invulnerable beings. However, despite philosophical concern that increased technological paternalism might be a reason for caution against certain medical interventions, transhumanists have not yet found a policymaking tool to tackle political-emotional obstacles to enhancement at the intersection of medicine with public will. I aim to show that nudge interventions improve public health in a way that answers this challenge, by influencing emotions that contribute to reasoned behaviour without unduly reducing autonomy or generating meaningless trends in behaviour.

Three substantial ethical problems facing Human Enhancement Technologies (HETs) in medicine

The transhumanist project aims at biomedical species-wide self-improvement and has been described as the controlled evolution, by humanity, of itself.²⁷ As a large-scale project that involves the enhancement of individuals, ethical transhumanism should respect individual patient autonomy and should be socially responsible to be called a success. This balance is also important in general medicine, where interventions like vaccination are generally understood to improve species wellbeing, though individuals might disagree or resist treatment in individual cases. Frameworks for medical ethics aim to pick out cases where intervention is acceptable, justifying certain instances of paternalism and denouncing others. In a similar way, transhumanist biomedical interventions divide public opinion despite evidence that they might be beneficial, and the benefits of any enforced enhancement might quickly be overshadowed by the injustice of impositions on freedom. A further concern for enhancement is the divisive effect of fierce disagreement on species-wide social, political, and ethical cohesion. I will argue neither curtailing individual freedom nor causing social division are inevitable for the transhumanist project, reimagining key transhumanist interventions within a libertarian-paternalist framework of public health responsibility.

Because transhumanism is an idea with a long history in many contexts, it is hard to define in terms of one goal or ideal. I will tend to refer to Bostrom's account of transhumanism as a project that aims to enhance the two core human capacities: lifespan and cognition.²⁸ Interventions that make human beings more resilient, flexible, strong, or resourceful; treatments that mean someone is able to jump higher, run faster, lift greater weights, to live longer, or reason more clearly, might be transhumanist interventions. There are a great many HETs which do not count as transhumanist, and about which it would be inappropriate to conclude that we were enhanced. For example, claiming 'the invention and widespread

use of the printing press was the moment when we became transhuman beings' would seem to miss the point. I will give more detail about the characteristics of uniquely transhumanist interventions, and explain why innovations like reading glasses, pocket calculators, and running shoes are also not transhumanist HETs in the short Primer chapter on transhumanist human enhancement that follows.

The benefits of transhumanist enhancement for the future of humanity are not yet fully clear and may never be completely understood, as more avenues for enhancement seem to become available all the time. Though the full extent of the impact of certain predicted enhancements on individual human bodies might not yet be clinically proven, for the purpose of this discussion I will take it that interventions at the point of their clinical use on human beings, will have been tested rigorously. I do not suppose that the invention of further HETs for medical use will mean the dissolution of medical trial and approval frameworks, or the rise of rogue, individual, enhancements. The question is not whether enhancement will become safe enough for human beings, but rather, once successful enhancement has been sufficiently documented in clinical trials, how we should go about integrating enhancement with existing medicine. Transhumanist enhancements may not yet have the unambiguous support of clinical evidence, but it seems all but inevitable that transhumanist enhancements of some kind will become safely viable for clinical use on human beings, in the near future.²⁹ Though I will not seek to defend this premise in the thesis, in the short primer to transhumanist enhancement, I hope to give an indication of the plausibility of transhumanism. The thesis will assume that enhancement options, once they have been clinically tested to meet currently acceptable standards and are considered safely viable, would need to be available to patients in a clinical setting, given similar restrictions and procedures for safe treatment that are required of general medicine. Although, as the transhumanist primer will indicate, enhancement also happens in independent and private laboratories, and is also a commercial enterprise, nonetheless using a clinical framework for ethical transhumanism productively narrows and directs philosophical discussion. Given this understanding of transhumanism as fundamentally medical and clinical, rather than as a private, hobbyist, or commercial project, the thesis will focus on the public health implications and the role of patient consent, individual responsibility, and social cohesion in public health.

Another challenge for transhumanist enhancement is that disruption to perceived or actual species unity, for example because of changes made to the core human capacities, might undermine social responsibility, undermine ethics, weaken the authority of justice and law,

and cause untold harm to many while advancing few.³⁰ The shared human condition that seems to be maintained by shared core capacities³¹ may be essential for maintenance and respectability of many ethical and social frameworks, including those explaining mutual responsibility and accountability. Disruption to social responsibility or the cooperative influence of empathy pose existential risks. In earlier political thought, J.S. Mill proposed that individuals with their own preferences and ways of life, each vying for their way to be allowed to flourish, were more likely to benefit the state than individuals pressed into conformity by the state.³² The mutual consent of citizens to some general understanding about what is normal, best, and appropriate, despite myriad individual differences in temperament and pleasures, relies on the understanding that there is a shared condition, despite individual differences. Blind justice can only be a realistic goal if all individuals must concede that they would all be equal candidates for the same misfortunes without the protection of their (contingent) characteristics resultant of birth or good fortune.³³

Transhumanism, therefore, represents an existential threat which might undermine the shared human condition that unites and strengthens ethical ties between individuals, despite other differences. Conversely, the authoritarian implementation of technological change to homogenise the enhancement of core capacities, also risks undermining individual responsibility and autonomy, particularly if it is imposed against the grain of public feeling. The forced homogenisation of core capacities poses an existential risk for those who fear the loss of an immutable core humanity and it raises serious ethical concerns about bodily autonomy and consent. Simultaneously, there is legitimate concern that the enforced implementation of enhancement to 'manage' individual traits pre-natally could promote the excessive homogenization of traits, for example to remove undesirable traits from the species altogether. This is a clear ethical challenge for enhancement, and one which means heritable gene editing of human embryos remains controversial and its implementation into general medicine does not seem likely to further global health goals, fairness, or justice.

Although at the extremes, both the extreme exacerbation of division between 'haves and have-nots' and the risk of excessive homogenization of traits, ethical problems seem comparable with those faced by any system of global distribution of resources, this is not enough to excuse them in an emerging system like transhumanism. Transhumanism, because its unique aim is controlled species-wide improvement, beyond existing natural constraints, should be fully accountable if enhancements were a deterioration of the status quo. This is no less true if transhumanism adopts entirely the same ethical approach taken by traditional medicine, because the onus is on the transhumanist to find improved

solutions, to protect against lowering the average quality of life for human beings, and because transhumanism already challenges standard medical frameworks. Because the transhumanist wish-list includes permanent hereditary changes to species traits, the onus is on the transhumanist to defend their particular project, to explain why it is more worthwhile to pursue transhumanism than to give preference to non-transhumanist improvements in general medicine. The social and political hurdles in the way of broad public acceptance of transhumanism are the result of longstanding perceptions of human health and flourishing. Longstanding intuitions about the human condition, bounded by core capacities, are hard to change, and defending the possibility of bringing about real, ethical, change in these intuitions will be the focus of the thesis.

Section One seeks to establish the groundwork for an argument based on the idea that the public perception of human enhancement, including general tendency of public feeling against authoritarian technocracy in matters of public health, poses a substantial risk to the success of ethical medical transhumanism. It might be the case that the problems resulting from negative perception of the means of transhumanism are more difficult to overcome than critical unease about its ends. The thesis argues that organised enforcement of group goals, such as intervention into public debate and about behaviour, to reduce unease and increase solidarity in matters of public health can sometimes be justified, particularly where novel technologies can change the choice architecture for the worse.

Secondly, that it is important to address these ethical hurdles in the social and political spheres at an appropriate and deliberate point (for example during processes of clinical trials), rather than leaving the promotion of transhumanism to the scientists and medics involved in improving the technology or to free market forces. In part, this is because ethical justification for the use of technology must be in conjunction with particular practices and outcomes (rather than judgements about the technology alone).³⁴ A parallel case can be seen in the increased difficulty for legislators over matters of 'smart cities' and face-recognition software, as the technology (though imperfect) has been in use for some time, is planned to become more integral to security in public places, but is not supported by public feeling or by its own legislation (for example about the secure use and storage of biometric information).

The species-wide application of any innovative use of medical technology depends heavily on political feeling and local social impact, and social debate is too important to be left uncoordinated. These are not secondary problems for transhumanist ethics but intrinsic to

its defence. It is less useful to discuss the global ethical status of a spade, if the biggest ethical concern is about the effect of spades on income and quality of life, than it would be to discuss the treatment of manual labourers by construction companies, or the distribution of spades in communities. Discussion of the risks and benefits early in the process should be important and an ethics of transhumanism must be predominantly the ethics governing social and political means to effect lasting change. To bring about enthusiastic participation in enhancement, this should be done earlier rather than later.

What the thesis will not do.

In this thesis, I will not attempt to prove that transhumanism has made accurate predictions about the future of human evolution, or that enhancements to the two core capacities are always good for the species on balance. Instead, both will be assumed, as well as the observation that aspects of human enhancement transhumanism are already underway, and that the success of some human enhancement techniques are now all but inevitable. Instead of defending whether transhumanism should happen, I will seek to identify why promoting transhumanist ideas could seem risky to fragile positive progress in personal responsibility for public health, for the success of global health standards, or for the health of the already vulnerable. In response to the challenge of these risks, I will argue in support of more organised and deliberate integration of transhumanism into medicine. I will argue in defence of transhumanism that it need not undermine or drastically reimagine personal responsibility, personal identity, or species unity in the process.

Moral enhancement is a further goal of transhumanism, which is usually discussed as part of the enhancement of cognitive capacities and is also deserving of philosophical attention. Better moral agents might behave more desirably but be less responsible for their actions. Loss of responsibility is certainly of concern to theories of ethical transhumanism that seek not to impede or undermine human autonomy in the process of improving the human condition. Concern about the risks and benefits of changes to core capacities might also be best understood following consideration of the way enhanced beings would feel about their condition, but this raises further argumentative problems.³⁵ So, chiefly for focus and the clarity of the thesis, I will restrict discussion to the enhancement of human core capacities at the genetic and cellular level that result in longer life and longer health-span. In brief, given fear and disgust over possible moral and cognitive enhancements, arguments in favour of pursuing moral enhancements are quite likely to be compatible with arguments presented

in favour of genetic enhancements targeting health, but will not be discussed in their own right.

The thesis will address concerns about adjusting the beliefs, desires, and emotions of the public that are detrimental to research development and public acceptance of transhumanism in medicine. The thesis argument is that a libertarian-paternalist or ‘nudge’ approach to governing public health responsibility is the best way to tackle several of the major ethical concerns over human enhancement, including fair distribution, justice, and the authority of healthcare providers to enhance.

In chapter 3, I will outline, briefly, the existing framework that allows public health policy to operate alongside advances in scientific medical understanding and available treatments, generating expert, activist, and layperson’s responsibilities. In order to carry out ethical policy making, it seems fair to suggest that government should be guided by the best available evidence and recommendations of experts within the healthcare domain. Best evidence and best advice about healthcare, which is regulated by a shifting paradigmatic understanding of medicine, is subject to change. I will not aim to defend the idea that paradigm shifts do happen in medical science, nor that a shifting paradigm is bound to change best practice and best advice available to governments. However, in the second dialogue of chapter 3, I will outline key disagreements in theories about how change happens in medicine as the groundwork for my later claims about global health policymaking as it has curtailed the progress of transhumanism so far. The idea of a global paradigm shift in health policy has been suggested and the growing field of public health has suggested innovations like a new Hippocratic Oath, which seems to offer new promise for enhancement as medicine.

Summary of thesis aims:

“We value our status as human beings especially highly, often more highly even than our happiness. This status centres on our being agents—deliberating, assessing, choosing and acting to make what we see as a good life for ourselves.”³⁶

Rogue enhancement experiments and the occurrence of individual biohacking cases do not constitute the species-wide transformation that transhumanism requires. The thesis argues that transhumanism would be much riskier if it existed only as a niche of independent research, occasionally provoking outrage with unregulated human trials but mostly

producing results about reduced senescence in mice. A socially responsible transhumanism should aim to bring about large-scale changes to public attitudes, to ageing, death, vulnerability, and the relationship between human agency and nature, to complement technological advances occurring in the sphere of research and development.

Socially responsible transhumanism, involving interventions to encourage and to normalise core transhumanist ideas should support both improvements in public health engagement and responsibility to public health in general, over individual interest. The thesis does not suggest that a personal preference for unenhanced human life, for gradual ageing and death is irrational, or necessarily identical with technological conservatism or a cognitive bias in every case. Rather that the public, who might have a variety of preferences about the interaction between governments, technology, and their own bodies, would behave very differently in the face of differently framed medical enhancement opportunities, given a nudge in support of their preferences, but not a push.

The approach defended by the thesis should also address a further concern, that expediting the general use of transhumanist enhancements to change human core capacities would risk alienation. Separate from the concern that enhancements would further divide the 'haves from have nots', critics of transhumanism have expressed concern that the rapid enhancement of adults could result in individuals becoming rapidly unrecognisable to themselves. One counter-position to transhumanism is that the current state of human biology is the best state for human bodies to be in, not because our lives are painless, always beautiful, or without difficulty, but because the limitations of biology allow us to be moral, rational, and autonomous. Hughes, for example, has responded to variations of bioconservativism critically, referring to its unwavering supporters as BioLuddites.³⁷

Rather than appealing to arguments from utility, for example transhumanism on the grounds that the increased health and core capacities of individuals would create a net benefit *despite* feelings of alienation and division, I suggest enhancement with the right kind of public support would be less likely to leave people untethered and alienated.

The implementation of any such large-scale public health project, aiming to reimagine deeply rooted norms about human wellbeing and lifespan, would involve top-down oversight to ensure fairness and protect against eugenics, as well as culturally appropriate branches developing practice to suit local populations. For transhumanism, as much as for any project involving medical research and distribution, cooperative top-down oversight by global regulators means treatment can be standardised, procedures can be improved and made

more efficient (with practitioners learning from each other) and is likely to reduce overall costs by coordinating efforts and sharing evidence. The thesis will approach the challenge of balancing top-down oversight and governance during a period of potential global paradigm shift, as one of balancing authority and the liberty of individuals. Though a kind of authoritarian control is required to bring such an ambitious project to fruition, this does not mean transhumanism automatically must involve forced homogenisation at the cost of individual autonomy or cultural identity, and answers to these kinds of challenges exist in other areas of public health policy.³⁸

I have so far indicated that the thesis is optimistic about the future of transhumanism in the public eye, and that the changes needed are possible. I will argue that transhumanism can be ethically implemented only if its goals and their progress are made public, and if the public mostly consider transhumanism to be acceptable in the ways it will be implemented. The thesis will assume that transhumanist enhancements are inevitable and most likely to be medical enhancements, which should be orchestrated by professionals in the medical sphere and so should involve regulation by governing bodies. However, it is not always the case that clinically approved medical procedures are acceptable to the public at the point where they can be legally distributed (consider birth control, vaccines, abortions). The tendency that people have to reject technological improvements on nature, at least initially, is of particular concern for human enhancement medicine, in part because of the reasons summarised above, that alienation, social division, and feelings of resentment and discomfort can follow. Critics describe these consequences as the inevitable result of a system that creates technological obsolescence, but where the obsolete product is the human body³⁹.

Social and political interventions accompanying medical human enhancement should aim to accelerate the pace of change in the following: public perceptions of ageing and old age; the importance of death as a part of meaningful life; the meaning of shared humanity. Successful human enhancement medicine should also actively mediate public reactions that arise against emerging enhancement technology in the way global public health organisations address global public feeling about birth control, vaccines, blood transfusions, and so on. The thesis will conclude that, although the success of transhumanism might not initially seem to require interventions on the same scale as those that might be justified by a sudden or explosive global public health crisis, its careful management is essential and relies in the same way on a stable ethical framework. Nudges are the best current solution to the problem of how to bring about global measures of oversight and support that are required during global health crises.

Ethical transhumanism must be transparent, should rely on informed consent (which can also mean tacit consent, given special conditions), and the products of enhancement should be available not only for private companies and private individuals but as part of the healthcare and health policy structures that are already in place to improve wellbeing, longevity, and to reduce the suffering caused by illness, which tends to be the focus of public health value frameworks. Correspondingly, activists who favour long-term transhumanist goals should first aim to make the public aware of, and happy about, the potential for general enhancements like longevity to influence their medical care for the better.

Human Enhancement and Health: Transhumanist Primer

In this chapter I describe the key goals shared by transhumanists, though transhumanism is a broad and varied approach to human enhancement and includes a wide range of technology, medical and non-medical projects. I explain why transhumanist human enhancement is not generally considered identical with technology driving advances in medical science. I explain why transhumanism is often controversial, is heavily criticised for promising panaceas that never materialise, and is sometimes equated with liberal eugenics.

Transhumanism is ideological. It is optimistic about the improvement of human bodies in order to reduce suffering, and draws criticism on the grounds that even its success would promote elitism and social division, technological authoritarianism, and homogeneity of human traits. I pick out two transhumanist goals, which will be the particular focus of the arguments of the thesis, the pursuit of robustness against ageing and disease, and extreme longevity. I suggest there is good reason to consider these the two goals that have always been at the heart of transhumanism, and that despite being the most lofty, these are also the projects of transhumanism which have the most in common with global public health endeavours, because of their focus on a general and widespread reduction in disease burden by means that can be inherited by the next generation.

What is human enhancement? What is Transhumanism?

To understand why ethical transhumanism would rely on an emotional response or the behaviour of the public, rather than the success of clinical trials and innovative invention, it is important first to understand what transhumanism involves and why it depends on more than clinical successes with laboratory mice, even if they have the potential to become successful human therapies.⁴⁰

In this primer chapter, I will define the key transhumanist ideas in one expert branch of human enhancement research and will explain how key ideas developed and supported by transhumanist and by immortalist activists are connected to gene editing technologies. I will introduce some of the more serious concerns about enhancement and critics of transhumanism. Having given a brief history of transhumanist ideas in philosophy, medicine,

and other disciplines, I will also say a little about recent transhumanist responses to criticism in the context of real human enhancement. Transhumanists suggest it is realistic that human beings might soon be medically enhanced, routinely, to live very long lives with elevated cognition.

The aim of this primer is to explain the work of the thesis which mostly discusses sociological challenges facing the introduction of enhancement into increasingly technological and individualised medicine. As, in the main body of the thesis, I will not attempt to argue that transhumanism or human enhancement are likely to happen, or that they are objectively the best possible direction for medicine, in this primer I will attempt to show why many people do not find transhumanist goals unlikely or undesirable. Although the thesis often describes enhancement as a kind of transhumanism that now seems inevitable given advances in medical research and clinical trials already underway, the Primer aims to establish that human enhancement medicine is transhumanism. I outline some key transhumanist objectives which, it is predicted, will very likely be met as the result of recent advances in medical science.⁴¹

This discussion divides human enhancement and transhumanism between what the ideas mean to experts, activists, and the layperson, to a distinction used in the literature on reproductive politics.⁴² Because reproductive politics deals with similar concerns to some of the criticisms of transhumanism, for example in the social and religious influence of belief on acceptable medical treatment and accessibility, the work of activists to promote and defend certain bodily autonomies and preferences, the risk of back street and off-target uses of reproductive technology, and the responses and behaviour of the layperson to all of the above, it is a useful distinction to help identify ethical approaches to transhumanism. In the literature on reproductive politics as in the existing work on transhumanist politics there is some overlap between activists and scientific experts, and some grey area also exists between activists and laypeople, but the general reference to three groups is useful, as each has different aims and different experiences of emerging medical technology in public health as in doctor-patient interactions.

Human enhancement: Expert research and expert criticism

Gene therapy was established in the 1970s. The aims of gene therapy are to identify, target, and improve genetic flaws that are the cause of disease in humans. By the mid-1990s, nine

research organisations were investigating ‘medically relevant’ human genes⁴³ with the aim of eradicating certain genetic causes of illness, and in recent years gene therapy has developed more rapidly. The introduction of new genetic material into the body, as one way to target genetic problems and change the characteristics of organisms, involves selectively ‘snipping out’ sequences of DNA and replacing them. The first successful application of CRISPR-Cas9 to target, remove, and replace an undesirable DNA sequence in human pre-implantation embryos was in 2015. In 2018, for the first time, human embryos edited in this way were implanted for pregnancy by a laboratory in Shenzhen, under He Jiankui.⁴⁴

Somatic cell genetic therapy alters the genetic material of the individual patient, as new DNA introduced to the body spreads through the somatic cells. However, introducing new DNA to only the somatic cells and not the gametes or sex cells means somatic changes made to the individual are not heritable. Early trials in somatic cell gene therapy have found some success, targeting single gene disorders like sickle cell anaemia, haemophilia, cystic fibrosis and immunodeficiencies, but do not yet have clinical approval for general use in medicine⁴⁵.

Germline therapy tends to be more controversial because changes made to the genes of an individual patient are heritable, *i.e.*, they may be passed down to any offspring of the individual patient. In many countries germline editing remains either explicitly illegal or is inaccessible because of more general ethical guidelines in place to reduce risk to patients of emerging and untested treatments. Ethical measures to restrict rogue human genetic testing, which tend to prevent the use of imprecise therapies with uncertain outcomes, also consider uncertainty about risks to the health of future generations.

CRISPR-Cas9 is one variation on CRISPR technology which uses non-viral means to introduce new genetic material to the body and was developed less than a decade ago.⁴⁶

In July 2019, faced with a World Health Organisation Report from Geneva which seemed hesitant or even alarmed, public enthusiasm about human genetic enhancement would have seemed foolhardy. WHO recommendations for laboratories working on CRISPR and possible applications of CRISPR-Cas9 technology for human enhancement were expedited, following the shock announcement from Shenzhen in the winter of 2018 that two babies were the first humans born with germline modifications. The early recommendations of the WHO, following the announcement were as follows:

1. That there should be established a WHO registry, to 'provide a more structured mechanism for collecting and curating details of planned and ongoing research relevant to its work'.
2. '[All] those conducting, or aware of research and development relevant to its mandate, in particular genome editing of human germline cells and embryos' should communicate with the Committee immediately. The Committee expressed a wish that unpublished existing research should be understood, underscoring the importance of WHO awareness of all existing work in the field.
3. Global health organisations should collect input from as wide a range of stakeholders as possible via 'an enhanced website; and targeted outreach to regional and country offices. Specifically, the Committee requested the Director General to engage WHO's regional and country offices and urge them to canvass societal views on human genome editing.' Cartoons and memes were mentioned as one possible means of communicating widely and across many languages.

A multidisciplinary WHO panel of professors in bioethics, risk assessment, genetics, technology, medicine, and law, addressed concerns that the apparent initial success in Shenzhen would lead to further human clinical trials. Recommendations were therefore focused on the state of the research at this key moment and intended to contain optimism about the research. With the recommendation of a central registry, and oversight of the WHO in trials underway, the report highlights the importance of slowing the advent of germline genetic modification into general medicine. This was justified in part by recognition of severe problems with the Shenzhen trial, both ethical and practical, further details of which have continued to come to light since the story initially broke in 2018.

Even before the birth of the children, He's trial did not adhere to the controls that are normally placed on agreements between clinicians and trial participants according to WHO guidelines for ethical human trials. The Shenzhen trial used stringent fines written into the experiment as a penalty if either family decided to pull out.

He intended to create a naturally occurring modification within a gene known as CCR5, which naturally reduces vulnerability to HIV. CCR5 is a relatively well understood gene that encodes the CCR5 protein. The CCR5 protein is present on white blood cells and acts as a 'gateway' for HIV entry into the cell; hence allowing for infection. There are naturally occurring

variables of the CCR5 gene within the European population, which He attempted to re-create.

Sperm-washing treatment was carried out which, along with the HIV-negative status of the mothers in the trial, rendered the modifications to the CCR5 gene of the embryos unnecessary by the standards of most practitioners. In terms of justification of the therapeutic use of germline editing, this has been described as moving the trial further away from therapy, and into the field of enhancement, as none of the children was at serious risk of being born with HIV, even before the intervention was made.⁴⁷

He's secrecy when acquiring funding to achieve these pregnancies was necessary because, although germline editing of pre-implantation embryos is a standard way to begin human clinical trials, the implantation of edited embryos for pregnancy is illegal in China and most other global jurisdictions. He's plans to announce the apparent success of the trial, with the birth of the children Lulu and Nana were uncovered before his announcement and were initially met with incredulity by the scientific community.

When the WHO met to confer again in 2019, the tone of their recommendations expressed severe reticence, and reinforced the norm that the technology, although promising, was not yet at the stage where human clinical trials were acceptable. It is still unclear to what extent the trial can even be taken as realistically indicative of the possibilities of germline editing for meaningful human enhancement. Targeting HIV resistance by other means which are less invasive, using non-heritable interventions, had already found success in recent years, in reputable trials, makes the true aims of the Shenzhen trial more questionable.

He's apparent success is predicted to have medical consequences for the two girls later in their lives, not only because germline modification is heritable (and so any children of theirs and subsequent generations might be affected) but because other areas of the genome than CCR5 are very likely to have been edited as well. These are known as off-target effects. As only 1% of the human genome is protein-coding, these off target effects are likely to take place in non-protein-coding DNA known as 'junk DNA'. However, many parts of this non-coding DNA are involved in the regulation of genes, and changes within these regulatory regions could have enormous consequences on cellular function and could ultimately lead to disease. Moreover, the ENCODE project has shown that 80% of the genome is used (transcribed), and therefore it is impossible to say with certainty that the off-target modifications won't have detrimental, or even enhancing effects on Lulu and Nana. The consequences of editing what seems to be 'junk DNA', surrounding the precise CCR5 area,

are as yet unknown and would likely also be heritable. Moreover, He did not make the modifications that he intended to make, and instead, novel changes have been introduced to the CCR5 gene – (i.e., not the naturally occurring variables of the CCR5 gene as mentioned before). Further, it has been shown that CCR5 is not the only gateway for HIV, so its choice in the Shenzhen trial was not fixed simply because HIV was the target. A further ethical question was raised regarding the consequence of CCR5 modification in China, as the variant is only naturally present in Europeans. Though the suggestion that the two children would be prevented, legally or medically, from having children of their own has been rejected as inhumane, the reproductive status of edited human beings had not been agreed or protected by global health organisations at the time of writing.

In the face of these events, the early caution of the WHO and public concern about the use of CRISPR on human subjects is natural. Nevertheless, the technology applied in these rogue trials has shown promise, and a significant improvement in terms of precision, time, and cost, on the previous Zinc Finger Proteins (ZFNs) and TALEN editing systems for making similar genome modifications. CRISPR is currently considered the easiest of the three systems to use, and its rate of off-target effects is considered variable (an improvement on ZFNs high rate of off-target effects, though there is a low rate of off-target effects seen in TALEN).

Improvements in the research into precise and efficient genome editing are likely to bring about improvements in general medicine and make truly individual personalized medicine a more realistic imminent paradigm shift. Somatic gene modification has already resulted in therapeutic medical interventions. In diagnosed patients, the treatment is normally *ex vivo*, meaning that cells are removed from a patient, modified, and then returned to the patient. CRISPR can be used to treat white blood cells extracted from HIV patients, and the CCR5 gene can be modified within the white blood cells. The CCR5 modification that occurred at Shenzhen could potentially be carried out on bone marrow derived stem cells, which may provide a long-term reconstruction of HIV resistant white blood cells. This type of therapy has already been found effectively to cure HIV.

The conduct of the researchers, the methodology of the Shenzhen trial, and the clinical treatment of the patients in the trial were criticised, as well as the legality of using CRISPR on humans. Concerns were expressed that the parents of the edited infants had signed contracts imposing strict financial penalties if they wished to withdraw from the test before the birth of the children. Financial penalties for withdrawal from human trials is widely unacceptable, given bioethical guidelines for clinical human research which prioritise patient

autonomy and self-determination. Particularly in the early stages of human trials, where it is generally preferred that participation should be entirely voluntary and should not take advantage of vulnerable individuals⁴⁸, the practice of fining withdrawal violates both patient autonomy and the principle of beneficence.

Controversy over the Shenzhen case only deepened when evidence seemed to suggest that the aim of the trial had not been to improve Lulu and Nana's resistance against HIV, as had been suggested. Instead, it was suggested by experts that the gene editing carried out on the embryos might also improve cognitive capacity⁴⁹. Suspicions that the team had concealed their true goals was reinforced because many argued that the application of still risky techniques to reduce the risk of HIV transmission between parent and child was unnecessary. Though HIV is still generally considered a relatively serious heritable disease, there are existing medical interventions that can be expected to reduce transmission rates from parent to child, particularly when only the male parent is a carrier of HIV.⁵⁰

Experts generally agreed of the Shenzhen trial that CRISPR was too imprecise for continued human trials in 2018, even if participants had been treated according to bioethical frameworks, which participants in the emerging case were not, and even if the aims of the study had been transparent. Despite consensus about the other advantages of CRISPR-Cas9 over previous technology, Zinc Fingers and TALEN, in the wake of the Shenzhen trial there were calls to identify and halt any other similar trials that might be underway. The WHO, in their early response to the Shenzhen trial, emphasised the need to create a global registry of gene editing with CRISPR, to keep track of the global progress of heritable gene editing, and to engage with as wide a range of stakeholders as possible⁵¹ so that further regulated moves toward human trials could happen in time.

Human enhancement: Activist perspective

Aubrey de Grey⁵² has responded to concerns about the risk of dehumanisation from making radical changes to human core capacities. De Grey identifies nothing intrinsically valuable about the 'natural' process of ageing and describes the negligible senescence as a clear good for health and wellbeing.⁵³

In Nick Bostrom's 'Fable of the Dragon Tyrant'⁵⁴, the position about reduction of suffering and the potential for a better human condition is presented in the most general terms possible. Bostrom uses the analogy of the dragon (old age and death) which terrorises a

society, and against which they must develop a technological weapon. In the Fable, the use of technology to remove a cause of suffering and death seems natural and common-sensible, and the critics of the anti-dragon technology seem short-sighted. In practice, the route to negligible senescence seems likely to involve so many individual enhancements to the bodies, capacities, and tendencies toward illness of individuals that arguments about eugenics and inequality cannot be so easily dismissed. Bostrom acknowledges this in his work,⁵⁵ but the simple rhetoric of the transhumanist against the obvious harm caused by death remains inescapably appealing.

Finally, De Grey⁵⁶ presents an account of normal ageing with predictions that are typical of the optimism of the transhumanist movement in general. De Grey explains human ageing in terms of seven causes, inter-cellular and intra-cellular, which impair the function of tissues, organs, and eventually lead to death of the organism. The account suggests that the list of causes of ageing, as it has remained unchanged for decades, despite substantial improvements in the understanding of senescence, is likely to be comprehensive. Solving the seven causes of ageing in practice would mean medical repair of age-related damage to human bodies that is accumulated during the normal activities of life. This level of cellular repair would mean the possibility of extreme longevity and negligible senescence: a life without ageing, perhaps entirely without death from decrepitude for anyone with access to treatment, which is tantamount to immortality.

But the transhumanist interest in gene therapy is not limited to improvements in healthcare (*i.e.*, the eradication of particular diseases like cancers or heart disease). Genetic testing, analysis, and editing can make both somatic (individual) and germline (heritable) interventions and have the apparent potential to eliminate genetic diseases and weaknesses of the human body, and to pre-emptively prevent damage accumulation in current patients and future generations. Technology that might be used to improve resistance against disease also seems likely to improve function and resilience of human bodies in general, reducing the number of possible causes of premature death, as well as reducing damage accumulation (*i.e.*, damage caused by cancer treatment, stroke, heart attack) during life. George Church is a particularly prominent advocate for genetic human enhancement. Since the early 1980s, Church's Harvard laboratory has carried out research into a wide range of gene editing applications for both disease-resistance and bodily enhancement. One popular page on the Church Lab. website is dedicated to a list of desirable outcomes of targeting named genetic sequences. The list gives examples of some genetic locations of disease as well as

interventions that seem more straightforwardly to be human enhancements and is often nicknamed the 'Transhumanist wish-list'.⁵⁷

Although many of the enhancements on the wish-list target single cell abnormalities in DNA sequences (or less beneficial variations that might occur), aim to treat recognised diseases like Alzheimer's, diabetes, cancer, and HIV, other items on the wish-list are not directly related to disease. Improvements on the list often come at some cost, to another capacity or characteristic, and are not straightforwardly likely to be useful to every human being, particularly in the non-disease cases. Enhancements to improve intellect, reduce vulnerability, and give greater insensitivity to pain, might be more universally appealing than increased bone density at the cost of buoyancy, but there are consequences of each, and no single intervention seems likely to bestow perfection or invulnerability. Instead, individual genetic enhancements improve the everyday capacities of individuals beyond the normal range, and for particular kinds of activities.

Definitions of enhancement do vary in the philosophical literature, with three main approaches finding different grey areas between treatment and enhancement, each arriving at different conclusions about what counts as treatment. Under Professional Domain accounts⁵⁸, medical treatments are any interventions that the professional standards of care endorse, while enhancements are any interventions that the professional medics declare to be beyond their purview. In the case of birth control in the 1960s and 70s, for example, general practice medics were initially uneasy about oral contraceptive pills becoming part of their domain, instead of remaining in the domain of family planning clinics.⁵⁹ Normal Function accounts⁶⁰ argue that to be healthy is to be able to do all that appropriately matched members of one's own species can do, for instance, what others of a similar age and the same gender can do. Health problems, diseases, or illnesses in need of treatment are characterized by a fall from the normal level of functional capacity expected for an individual in comparable range. For Disease Based accounts⁶¹, treatments are interventions that address the health problems created by diagnosable diseases and disabilities. Enhancements, on the other hand, are interventions aimed at healthy systems and traits.

Each approach has its shortcomings, but disease-based and normal function accounts are more often used to discuss the divide between transhumanist enhancements and enhancement-like general medicine, because of the normal meaning of the word enhancement, which is suggestive of augmentation rather than cure of the human condition. However, as in the case of the transhuman wish-list, some interventions target recognised

disease states while others do not. Critics of the disease-based approach argue that there is a strong historical precedent for pathologizing and de-pathologizing certain conditions like gender dysphoria with significant consequences for the treatment of individuals.⁶² This leaves disease-based accounts open to criticism for being still too ad-hoc to form the basis for making sure medicine does not slide into enhancement.⁶³ This is important because, whichever account is taken to give the best explanation, enhancements tend to be more controversial than treatments.⁶⁴

Longstanding concerns about the pursuit of enhancements like those on the wish-list include the fear of eugenics, concerns about increasing homogeneity of the species, and forced enhancement. Concerns about enhancement leading to eugenics include both typical eugenics and what is known as 'liberal eugenics', where certain individuals are encouraged to have more children than others, on the basis of characteristics, social status, or their beliefs and affiliations. Concerns about homogeneity, about losing diversity, come from concern about eugenics. However instead of removing certain individual human beings from the population able to pass on their genes, it is possible that genetic enhancements might mean altering the heritable characteristics of a generation to some desirable standard. Losing trait diversity, *e.g.*, all children born to an edited generation will now have blonde hair, or stand over 6 feet tall in adulthood, seems inherently undesirable. Concerns over forced enhancement predict that as certain traits and capacities become valued over others, individuals would be under more pressure to undergo treatment or enhance their offspring, or else risk social disadvantage, discrimination, and dehumanisation. Subsequently, the concern is that sufficient pressure would be tantamount to forced medical treatment, and incompatible with ethical good practice in the case of most kinds of medical interventions.

Church has responded to concerns about the sudden and widespread effects of facilitating genetic changes to the species, risking homogeneity across generations, about characteristics that may later be de-pathologized.⁶⁵ Church has suggested that some technology already in circulation seem to count as transhumanism, if taken out of context, as they make substantial changes to the activities and quality of human life with similar impact on diversity of behaviour and capabilities. The change in capabilities that became available to most human beings when, for example, fire and compound sentences became commonly used, were significant and entirely transformative, and constituted kinds of technological change to the human condition from which there seems to be no return. One problem with reference to examples like fire and language is that they are too distant and abstract to be useful or meaningful, though we might fancifully imagine the effect of such

changes on human lives. It is easier to comment, however and the products of commentary are more rigorous about the impact of more recent technological advances. The increasingly efficient and inexpensive combination of mobile technology with a wireless internet connection allows for such widespread incorporation of 'Smart Phones' into all parts of home and professional life. The use of mobile phones has changed in the last three decades, almost beyond recognition in comparison with the first instances of the technology in the 1980s, and ever-changing applications have also changed behaviour and cognition in normal life.⁶⁶ Changes in normal behaviour and normal understanding of the behaviour of human beings, results from the widespread use of emerging technology, and this change can happen suddenly. In some cases, fast technology-driven changes to human life and understanding of human bodies have led to some scholars re-formulating and reconsidering philosophical thought experiments about the conditions needed to create extended minds.⁶⁷ The sudden availability of technology like smart phones is a kind of enhancement which became available to generations of people at once, with widespread effects on behaviour and capacities. These widespread effects happened quickly and have been left almost unchecked, without the gradual acclimatisation of norms and behaviour that would be more likely with a genetic enhancement which would take decades to move through generations of people.⁶⁸

A more limited Transhumanism focuses on permanent body modification which is carried out for enhancement of the human core capacities. This limited definition is more useful for philosophical discussion because it creates a focus on the next likely transhumanist enhancements that will become available to human beings, without overlooking how close to transhumanism some accepted medical technologies might seem. Emerging medical technologies are more deserving of the transhumanist label than emerging technologies outside the medical sphere, because of their power to target core capacities like lifespan and cognition directly. Whether there is good enough reason to differentiate sharply between social and medical enhancements that improve human life, such that one is more obviously transhumanism, is not always clear simply because they can each make life easier, longer, and healthier for human users.

David Eagleman has shown great public optimism for the (near) future of human brain enhancement, rooted in technological augmentation to allow better and more sensation of the immediate environment. Eagleman himself is enhanced⁶⁹ and has predicted more widespread use of enhancement technology to improve human interactions with environments (heat sensors, data stream trackers which could be used to 'feel' the stock

market or company productivity, electromagnetism, and echolocation, as a few examples) in the next five years.

Transhumanist Human Enhancement Summary

In this chapter, I have limited the definition of transhumanism in this thesis to *the enhancement of human bodies such that the core capacities either of individuals, or of individuals and their potential offspring, are improved outside the normal range*. This definition excludes organ donation, birth control, vaccines, and other medical interventions which return the core capacities of individuals to the normal range, even though they may improve the health of some individuals beyond a range they had experienced before.

As a further example of how transhumanism highlights difficulties for existing morality, consider ethical arguments about the ‘non-medical’ or off-prescription use of cognitive enhancement drugs for study purposes. It is unclear whether certain drug-taking behaviour by individuals that physically affects only their own bodies is acceptable or should be legal. One longstanding position from the philosophy of medicine aims to prioritise the autonomy of the patient, whereas public health ethics take a more critical view of individual bodily autonomy, where individuals must be considered part of a society: Their health is part of a shared public health. Unease about enhancements becoming part of healthcare focus on concerns that artificial motivation, energy, capacity for concentration might lead to society denying or undermining the accomplishments of naturally endowed students, or athletes. In favour of enhancement, transhumanists might counter the *prima facie* intuition that study drugs are not acceptable by presenting a comparison with existing standards that are applied to caffeine use in the workplace, for example. Comparison with the use of an already socially acceptable stimulant shows how unreasonable is the perception that the drug alone is *responsible* for the work of those who, for example, drink coffee. Comparisons like this, with apparently everyday substances that change behaviour and capacity for behaviour can also help make sense of the argument that there should be no meaningful social division between those who use study-drugs and those who do not. This is to oversimplify controversies arising from off-prescription cognitive enhancement and ethical questions about performance enhancement in sport but does demonstrate the importance of appropriate framing for enhancements. Referring to off-prescription drug use both as a kind of enhancement of core capacities and akin to everyday coffee-drinking behaviour, might shed new light on the role of core capacities in meaningful activity. Transhumanism tends to be in favour of human

enhancement, surmising that it is “morally permissible (and sometimes obligatory) to use biomedical means to modulate or select certain biological traits in order to increase people’s welfare, even when there is no pathology to be treated or prevented.” Further than this, many transhumanists maintain that ageing is pathological and can be cured. This difference in opinion is important to understand to make sense of the challenges for certain pro-enhancement arguments entering medical practice. In the second dialogue, I will defend the idea that transhumanism could become a successful medical paradigm despite significant differences of opinion about death and ageing in the field of general medicine, because of the salience of ageing as a shared human experience to public health goals.

This approach to defining transhumanist enhancement relies on accepting a normal range of human core capacities. The benefit of accepting a normal range of core capacities, rather than a fixed set of essential characteristics of human beings is that it is appropriately flexible to the real experience of variation in human health. Theoretically, this approach still allows for the possibility that some human beings might themselves fall naturally outside of the normal range because of genetics or nurture. For example, it is not within the normal range of cognitive capacities to be able to recall every waking moment since birth with clarity and without deterioration, and any person with the ability would not be cognitively within the normal range. Although encountering one person who is able to recall memories in this unusual way does not make the ability normal, it does suggest the capacity is in some way a human capacity or could become a normal part of the human range if more individuals had the capacity. Likewise, the normal range of human lifespans currently extends to around 120 years maximum, but a person found alive at 140 years of age would not immediately be inhuman simply because of their age. After a human lifetime within the existing normal human range, with the experience of ageing that their 140 years would cause, there would be no reason to suspect the individual was a different kind of organism altogether. These cases, because they are extremely unlikely, need not undermine the idea or use of a normal range, and need not lead to declassification of certain individuals as human beings. It makes the most sense, when discussing medical human enhancement, to argue that the deliberate enhancement of an individual, resulting in significant increases to one individual’s memory or lifespan also would not change the normal range of human core capacities. Until the effects could be reliably achieved, could reasonably be described as normal, or were accessible to populations at large, the fringe enhancement of individuals beyond the normal range would not be a successful transhumanism.

Three Key Dialogues from the Philosophical Literature

Introduction

There are many reasons philosophers should find transhumanism interesting. In this chapter, I draw out connections between transhumanist issues and existing philosophical problems. I frame transhumanism as neither an answer to any of these problems in totality, nor as simply an interesting example of any of these problems, but rather as the next step in an ethical project of improving the position of human self-understanding as it affects problems like anthropocentrism, the uniqueness of human reason, and moral imperatives that seem to affect only our species' activities and development:

First, transhumanism challenges fixed ideas of human nature by promoting and predicting human self-evolution.

Second, transhumanism questions whether shared biological characteristics drive social cohesion, empathy, cooperation, and morality.

Third, and particularly because it predicts extreme longevity and advanced cognition, transhumanism offers a revised account of what counts as meaningful human activity. Transhumanism suggests human lives can be meaningful even when they stray from the apparently natural conditions of life, like fierce competition, scarcity, hardship, and death.

Transhumanists therefore question certain existing accounts of personal autonomy and responsibility, challenging notions of personal responsibility, and suggesting stronger ties between human activities now and future human endeavours. One potential consequence is that individuals should take more responsibility for decisions that affect future generations, as transhumanists tend to suggest we can act collectively as a species, and for the benefit of the future of the species.

Transhumanism also appears to challenge a certain kind of anthropocentrism which treats the current human perspective and approach to life as fundamentally better and more important than any alternative. Transhumanism also tends to promote greater awareness of the negative human impact on the long-term condition of the planet, and the damage we do to other species. Transhumanists hope to bring about technologies that might forever change the scope of our responsibility to do the right thing for future generations, for

example through heritable genetic changes, in part by recognising how damaging and flawed human activity is.

Meaningful life, autonomy, and responsibility for change are each the subject of longstanding philosophical investigations in the philosophy of medicine and in moral philosophy beyond transhumanism, but transhumanism presents an opportunity to consider our species and its activities from a different perspective. Sometimes, transhumanism is used as an example of unconventional but undeniably human activity, to establish the flexibility of a certain approach to defining autonomy, responsibility, humanity, or personal identity.

Consider, for example, the trend in the philosophical literature on meaningful life toward a more piecemeal understanding of what the 'bearer' of meaning can be. A gradual shift away from conservatism or essentialism about human purpose or function in ethics has also weakened the whole-life views of meaningful life, and those based on singular lifelong pursuits or 'higher' ambitions alone.⁷⁰ Any approach to meaningful life that acknowledges and welcomes the plasticity of the human condition is more likely to result in accounts of meaningful activity where the bearer of meaning can be an instance or event instead of the completion of one strong narrative arc.⁷¹

In another kind of philosophical argument, transhumanism might be presented as a hypothetical, or as a thought experiment that seems to undermine certain metaphysical answers to longstanding and troublesome questions like 'what is personal identity?' or 'what is life?'. Perhaps the best known contemporary example is Hilary Putnam's 'brain in a vat' scenario.⁷² Even given this use of transhumanism as hypothetical cases to define philosophical 'grey areas', the increasing likelihood of real transhumanism, for example the increasing potential to build a lifelike world-simulation vat, presents problems that existing frameworks for practical ethics may not anticipate.

A more sociological example of the latter kind of contribution made by transhumanist philosophy can be found in arguments about the role of boredom in maintaining a strong sense of purpose or identity. If meaningful life relies on the whole or complete life, immortality seems to pose a problem, and if meaning relies on completing a single narrative arc, then immortality decreases the likelihood that anyone could find meaning. A prominent argument about the loss of meaning resulting from the inevitability of boredom for immortal beings was introduced by Bernard Williams and did not intend to discuss real immortal beings, but presented the problems of a fictional immortal as an example to anchor arguments about death and meaningful life.⁷³ More recent responses to Williams' Necessary

Boredom Thesis however, work on an updated assumption that some form of immortality is an increasingly real possibility, and contemporary counterarguments about the role of boredom for immortals are correspondingly different from Williams' proposed case from fiction. I will revisit arguments about the role of boredom in the life of an enhanced person in chapter 6.

The second dialogue must explain how scientific shifts in understanding should come about, and how corresponding shifts in the public comprehension of best available explanations arise from science. Medicine, as an evolving scientific discipline, also relies on a trusting relationship with the public, and on a level of public health education. Changing the basic assumptions underlying progress in a discipline, and the finer-grained elements of good practice within that discipline, also depend in part on interpretation and implementation of accumulated convention and practice. Conventions often change gradually as the result of innate resistance biases; limitations resulting from practical contingencies of funding, location, workforce; and the need for rigorous evidence accumulation, review, and dissemination. The idea of paradigmatic shifts tends to be informally understood in terms of long periods of stability broken by sudden leaps. This is a simplification, and paradigmatic science involves constant movement during the periods of comparative stability as well. A finer point of disagreement about what constitutes a paradigm shift within medicine concerns what would constitute the collapse of general medicine into enhancement medicine. This has bearing on ethical judgements about practice, governed by accumulated convention under one paradigm, before it becomes appropriate to shift to the next. The contributing role played by bias (particularly internal bias and the status quo bias) in approaches to public health and medical paradigm shift will be identified. Depending on the significance or value placed on maintaining gradually occurring paradigmatic changes (as opposed to the preference for an expedited progress, perhaps by nudging) in the medical sciences, ethical transhumanism would be a different phenomenon.⁷⁴

The third dialogue will address a difference in theories about the role of emotion in reason. While it is increasingly common that accounts of reason do not deny the involvement of emotion, the implications for behavioural policy vary. For rationalists, emotions cloud reason and should not inform theories about rational normativity, for subjectivists the emotions, although they do not aid reason, should be included in proposed explanations of rational normativity. Most promising, though, is the cognitive position, asserting that emotions are necessary for reason, and should be considered in any evaluation of rationality. Whether the emotions are considered as reduced to beliefs or desires, driving intentional action; whether

they only seem to behave like beliefs or desires; or else carry beliefs or desires with them, which means a dual explanation of behaviour caused also by desires, emotions guide decisions and should be considered important for evaluation (like judging risk).

First Dialogue: Autonomy and Behavioural Insights into Free Action

The first dialogue will address a disagreement in the philosophy of autonomous action, concerning the extent to which external pressure undermines agent autonomy. Autonomy is central to ethical general medicine as one of its pillars of good practice. Patient autonomy is important for patient consent, which is a legal necessity in modern medicine in much of the world. I will first summarise the disagreement about acceptable external pressure on behaviour, before applying a practical account of patient autonomy to difficult medical cases, to test the limits of informed consent. Toward the end of the dialogue, I will describe the account of autonomy that is most fit for use in emerging transhumanist ethics.

Behavioural insights or 'nudge' theory is an approach to behaviour management which aims to intervene on the behaviour of individuals without recourse to overtly paternalist or authoritarian control or restriction of choice.

"Nudge. A nudges B when A makes it more likely that B will ϕ , primarily by triggering B's shallow cognitive processes, while A's influence preserves B's choice-set and is substantially noncontrolling (i.e., preserves B's freedom of choice)."⁷⁵

Nudges make it more likely that an individual will behave a certain way, and a nudge can be either overt or covert, sometimes working best when they are not obvious to the nudged party. Sometimes, nudges can be fully apparent to the nudged party and still fully effective, but this does not always mean the action of the nudge is clear, or that the nudged party fully understands how they have been influenced. In some cases, nudges act on or against implicit biases, and are most effective when the nudged party is unaware of what their initial bias might have been.⁷⁶

Definitive examples given by nudge theorists tend to be those where the physical environment is subtly reorganised as in two famous examples. First, in the school cafeteria example,⁷⁷ the server must decide where to place healthy and unhealthy snack items, knowing that certain placements make it more likely that more children will choose the healthier option. The example is archetypical of nudge theory because it influences

behaviour without removing any option from the children, or making any option harder to reach, and is interesting for three reasons

- i. The children are not prevented from choosing freely, even if their choice is the less healthy snack option. They have not been redirected to choose between healthy options, nor given fewer choices.
- ii. The children are not aware of the nudge. There need be no signs in the café suggesting healthy eating, and no information about the options available for the environmental change to be effective.
- iii. The server has no choice but to position the snacks in *some* arrangement in the cafeteria. If both healthy and unhealthy options are to be left available, they must both be somewhere, and truly neutral environmental design is unlikely by chance.

In the first case, it is important that a nudge should not be able to, nor aim to prevent undesirable behaviour by force, nor to restrict individuals by limiting their available choices. The choice of a chocolate bar could be, in the relevant sense, prevented if there was no chocolate in the cafeteria, or in any cafeteria, but this would not be a nudge. The choice of a chocolate bar could be restricted if the price of chocolate were increased to make it too expensive for many or most children. This would be a kind of social engineering with more nudge qualities than an outright ban, but still without the characteristic openness of a nudge, because certain people would be prohibited from enjoying chocolate, more than other people. In this simple café example, identifying interventions that restrict and limit freedom of choice is relatively straightforward.

However, in the wider social use of nudges, the impact of interventions is harder to predict, and less simple, and some legitimate nudges do also include financial incentives or targeted schemes that have disproportionate impact on marginalised groups. It is generally less controversial to nudge without using financial incentives, or example, in the UK Government's early Behavioural Insights Report, MINDSPACE⁷⁸, incentives in general were a key part of the framework for behavioural insights work, with financial incentives as one possible instrument for making behavioural changes. One example would be taxes on the sale of tobacco or the import of certain products to make them less attractive but still legal. However, in Thaler and Sunstein's introduction to the concept of a nudge, financial incentives and penalties are disallowed as they 'place too great a limit on freedom of choice because of the price increase and cost burden'.⁷⁹

Since financial nudges place financial barriers in front of some behaviour for certain groups (and, in particular for more vulnerable, less privileged groups) they do not seem akin to less coercive behavioural insights measures, though they also give an insight into reasons the public might act or fail to act.⁸⁰ In the later UK Government Behavioural Insights Report, EAST,⁸¹ financial incentives were no longer a suggested part of the nudge framework.

Second, the children do not know about how or why their cafeteria space has been engineered, or how the nudge effect works on them. They may know, however, about the nudge salient concerns of healthy eating, even if they would knowingly choose unhealthily for themselves. It is assumed that the nudge does not act in a way that is alien from what the children would understand about their own best interest, and that it promotes the values that are expected of them by the adults in charge. In this case, health is the key concern, rather than fruit eating, and the nudge encourages the children to behave in the way most likely to result in their improved health.

Third, the server cannot fully avoid the responsibility of organising the environment, once they are aware of the effect of snacks on health. If the school, the server, or the adults in charge of caring for the children during their time at school want to make a positive impact on health without authoritarian control of their behaviour, awareness of the impact of snack placement on choice is useful. If the school refused to take an active interest in healthy choices at all, the environment would still impact behaviour, but in untold ways, so *some* level of responsibility cannot be avoided by those with authority.

So, all the children must make a choice, and they may have different levels of awareness about the fact that there is a choice, or that it is a choice relevant to their health. The children, as they are children, may not yet feel strongly about their own health, and because they are children it seems more appropriate for their environments to be organised to promote valuable goods like healthy eating. The nudge benefits the health of the group, regardless of individual variations in knowledge about healthy eating, feelings of personal resistance against the healthy eating values coming from their school, or willingness to act on what is known to be healthy in their own interest.

In the second example the aims of the nudge are similar, but the target audience is broader, and the setting is more public. Where a staircase is next to a lift or escalator, painting the staircase in bright and appealing colours, or adding a pattern or a game to the stairs, means more people are likely to use the stairs than to ascend using the less labour-intensive option. In Stockholm, a viral marketing campaign by the car manufacturer Volkswagen led to a

reported 66% increase in stair-use, in the Metro system next to an escalator,⁸² demonstrating that an appeal to a sense of game-playing can be a powerful behavioural tool. In both the school cafeteria and the painted staircase examples, changes are made to what Thaler and Sunstein call 'choice architecture'. The differences between this nudge and the school cafeteria are important to note, however:

- iv. The target audience of the staircase is adults of all ages, and the children accompanying them on the Metro. The audience is broad, and the setting is public. The staircase is not in an overtly educational environment, and does not target only children, who are not yet considered fully competent choosers.
- v. The strangeness of the painted staircase nudge, particularly in the viral marketing musical stairs instance, makes it more visible than fruit placed at the counter in a cafeteria. The nudge might be even more overt than just eye-catching bright colours, involving a stair-climbing game or reward of some kind, suggesting immediate incentives for taking the healthier option (i.e. musical notes sound when the action is completed, 'calories burned' painted on the stairs up,⁸³ or perhaps a patterned walkway to follow, offering a feeling of 'completing the challenge'⁸⁴). Like the school cafeteria, though, the nudge does not need to explain itself or present its reasons in order to have an impact on behaviour.
- vi. Like the school cafeteria, those responsible for designing the environment cannot help but nudge the environment users one way or another. However, those nudged by the painted staircase may have more options to avoid the particular nudge environment than do school children seeking their lunch. The location of a nudge does change its reach and its significance, and some nudges could change the choice architecture in environments that adults could not avoid, or that some marginalised groups in particular would find it hard to avoid.⁸⁵ This would make the nudge more like the school cafeteria, and would require strong justification, given the difference between organising and controlling the behaviour of schoolchildren in a learning environment, and treating adults the same way. The autonomy of the target group and the freedom of adult groups to avoid certain nudges completely are factors to consider in evaluating each nudge case.

Both examples were intended to improve the health of groups by encouraging behaviour that would generally be considered healthy (valuable) in our society. The restriction to freedom, and the potential negative impact of the nudges were minimal in each case, and the values involved were commonly held. Nudging has been the focus of a policy group advising the government of the United Kingdom since 2009, and their first report on nudges identified many more approaches to behaviour management than changes in salience, which best describe both examples above.

The report also discussed the impact of the messenger, incentives, defaults, and more factors that change the presentation of choice: the choice architecture. Within this framework, behavioural insights techniques varied, and some nudges were more effective or easier to justify in certain situations over others. Not all nudges are effective in all situations, and nudges can be overused and misused to the detriment of public responsibility or beneficial behaviour. A nudge that is acceptable in one setting might be too invasive or out of touch in another. In some settings, the 'gamification' of positive behaviour might seem inappropriate given the seriousness of the content of decisions, but sometimes introducing fun even into serious activities like breaking addiction is appropriate, harmless, and effective. There are often good reasons not to incentivise desirable behaviour with financial rewards. Beyond the difficulty that financial incentives can work with less force on the wealthy, and with more force on people without financial security, financial incentives are often less effective than other kinds of nudge. A time-limited programme to incentivise desirable behaviour with financial rewards might risk reducing the likelihood that anyone would continue the good behaviour in the absence of the nudge, once the expectation of compensation is created. When nudges are time-limited and do not change the choice architecture in a sustainable way, for example by creating habits or commitments that last, financial incentives risk temporarily increasing the cost of existing good behaviour to governing bodies, without creating lasting change.⁸⁶

Nudge appropriateness depends on various different ways people judge the choices they make. In some settings, it is more important that individuals feel they are 'seen' while choosing. In other settings, it is important that governing bodies intervene covertly to see results, whilst being fully transparent about their policy in general. Given sufficient justification for pursuing the underlying value behind the policy, (*i.e.*, healthy eating or the need for organs for transplant), it is not always necessary that attention should be drawn to the significance of individual choices, or the fact that nudges are taking place. In the case of healthier diets for school children, and more frequent staircase use by the general public,

the relevance of the particular behaviour to the core value, public health, is relatively uncontroversial and relatively well supported by medical evidence, so this clear causal explanation between the behaviour (healthy eating/ exercise) and valued state (health) makes the intervention easier to defend. It is less important to explain, as in the cafeteria and staircase examples, what is happening or why certain behaviours are valued, whereas in other examples it might be important to draw attention to the choice and the benefit of a particular choice, as part of the transparency of the intervention. More will be said about the justification of nudges and their relationship with tacit consent in Section Two.⁸⁷

What does it take to undermine autonomy?

In this section, I will outline an account of personal autonomy, explaining why it is difficult to find satisfactory necessary and sufficient conditions to define autonomy clearly. Assuming that there is personal autonomy, despite these challenges of finding a comprehensive definition, and assuming people act autonomously in their everyday lives, as it seems like they do, I will first discuss the challenges that seem to undermine everyday autonomy. In the second section, I will turn to disagreements about when a restriction of that personal autonomy can be justified in matters of public health and medical practice. Whether nudged behaviour can be called autonomous, and whether strategies of government involving nudging would unjustifiably interrupt personal autonomy will be important to settle in defence of transhumanism. Transhumanism seems to involve individuals taking more responsibility for their own core capacities as human beings, as well as necessitating a level of authoritarian control to help change species-wide wellbeing. So, transhumanism is personal, social, *and* political, and the ethics of transhumanism require a tailored approach to this political problem of autonomy.

The political difficulty of protecting individual liberty under democratic government is a longstanding matter of philosophical discussion. Philosophical descendants of J.S. Mill take seriously his concern about the need for self-protection against the 'moral coercion of public opinion,'⁸⁸ which was initially a reaction to increasingly democratic, and utilitarian, systems of government. In a system where the good of the public is a concern for the individual and given a certain scepticism about any *a priori* insight into objective truths, Mill recommended general openness of thought and freedom of action. Openness to the truth of empirical claims was intended to promote liberty and progress, supposing it more harmful to redact or suppress some true assertions than to accept and consider some false assertions. Mill's

preference for openness, meaning greater freedom of thought and discussion, he believed would lead to faster empirical progress, where communication and examination of information would happen more widely, also meaning truth would proliferate more quickly. Openness and freedom of thought would also allow what Mill described as the ‘experiments of living’⁸⁹ to go on, mostly unrestricted by law. This openness and freedom of thought would lead to greater individual variation and greater diversity of lifestyles, not to be restricted unless to stop harm coming to others. The competent adult, with the capacity to rationally reflect on their desires and beliefs, Mill supposed, was far better able to self-govern and to form their own life than any paternalist government. The political should therefore refrain from action to restrict liberty, and the only just use of political power⁹⁰ “over any member of a civilized community, against his will, is to prevent harm to others.”⁹¹

The influence of Freudian commentary on the psyche later saw philosophical accounts of human behaviour shaped by the recognition of latent, primal, and subconscious desires and motivations.⁹² The suggestion that human beings were not fully self-aware and critically reflective appeared to undermine the rational, competent self-governing individual in Mill’s open marketplace of empirical progress. In the wake of psychoanalytical theory and later neurological experiments on the causes of human reason, modern accounts of personal autonomy still retain the libertarian preference for diversity and freedom of life. However, these accounts are more critical of the capacity of individuals to choose the best for themselves following rational deliberation. The exercise of authority over individuals, to ensure better outcomes for public health, education, and scientific progress, seems a natural consequence of the judgement that individuals are part-driven by biological and deterministic subconscious processes, and may not always freely choose what is best.

In the absence of direct coercion or compulsion, an agent living within social conventions and the law might feel they are free. The experience of feeling motivated by personal reasons, of assessing situations, and of knowingly consulting one’s own beliefs and desires before making a choice might seem to support this. However, this feeling of having freedom might be insufficient in light of evidence from psychological and neurological studies suggesting subconscious factors govern behaviour to a greater extent than agents feel, and that external forces determine action. It is increasingly common for the literature on personal autonomy to refer to studies that demonstrate the prevalence of unconscious factors seemingly hidden to the agent, which cause action. An agent might explain the reasons for their behaviour but be prone to bias and confirmation effects they do not recognise. Additionally, the kind of moral coercion from society described by Mill need not

be overt or consciously recognised by the agent for it to influence their behaviour. It therefore seems possible that an agent might report choosing A instead of B for reasons of personal preference, that they might be correct about their preferences, *and* that they might be wrong about the direct causes of their behaviour. In misjudging their freedom to have acted otherwise, the agent's feeling of freedom might not mean they are free. The disparity between agent self-report about the feeling of freedom, and the scientific suggestion that behaviour is guided by determined causes seems to suggest that there is more to autonomy than the mere feeling of self-control or freedom from compulsion.

If it were a necessary condition for autonomous action that it should occur without or despite all outside influences and biases, we would need to accept that no action is autonomous. This conclusion is counterintuitive, particularly in situations where agents seem responsible or blameworthy for their behaviour. An account of autonomy as action without subconscious bias or outside influence therefore seems too demanding to apply to everyday notions of acting autonomously, regardless of whether arguments about determination follow. Evidence suggests that action, even when it feels autonomous to the agent and undertaken for reasons of desire, motivation by reasons, or to pursue preferences, is very often the result of social causes, including those operating beneath self-awareness or beyond conscious control. It is unlikely that such causes of action would be reported in the explanation given by agents of the reasons for their own behaviour, even if agents were aware of the theory. In the face of strong evidence that mitigating circumstances, environment, conditioning, and subconscious biases play a part in human reason, some philosophical accounts of autonomy suggest behaving truly autonomously is compatible with the influence of determinism. However, within compatibilist accounts a further disagreement remains, about the acceptable limit of external influences on autonomous, self-controlled, actions.

According to a compatibilist account of personal autonomy, autonomous action and determinants beyond agent control are not mutually exclusive. Compatibilism still requires that desires and beliefs, as well as the ability to reflect critically on desires and beliefs are necessary conditions for autonomy. Autonomy is possible, even in the presence of determinants that seem to cause action, and which could be used to give a full explanation. However, even assuming everyday notions are sufficient, and even defining autonomy loosely, autonomy suffers if the agent can express no beliefs or desires relevant to their action.

In situations where new information is presented, or where persuasive discussion influences the individual, autonomy could remain. Even if the explanation were that external factors had completely reshaped desires and beliefs, disrupting reasons, an agent seems autonomous. Persuasive influences might motivate an agent to choose differently than they otherwise would, leading to different reflection on existing beliefs or desires, and this seems to match the everyday phenomenon of autonomously changing one's mind. Another important feature of autonomous agents, then, appears to be the capacity for critical reflection, which is common to most human agents, and allows the synthesis of beliefs and desires in the face of new information and experiences. This capacity for reflection helps make sense of the apparently high incidence of personal autonomy in adults, even in those who are easily persuaded or change their minds often. Autonomy remains in such cases, despite strong evidence for the ever-present influence of external factors which would explain the acquisition of beliefs and desires, and the causes of action. This also suggests that individual liberty to self-govern, conditional on causing no harm to others, and given a capacity for self-reflection might still produce the best overall outcomes, as in Mill.

Given an account of autonomy incorporating the problem of apparent determinism, even if an agent reports preferences, desires, and rational deliberation as causes, actions that can also be explained solely by the effect of external determinants might appear not to be properly autonomous. To return to the conditions of liberty in democracy imagined by Mill, the autonomous person would not be *necessary* in order to explain progress or diversity present in the 'experiments of life'. It is important, therefore, to be able to distinguish between determination and coercion. Habermas', arguing against human genetic engineering rests on two premises; 1) that protection of anthropological identity is important if we wish to retain moral dignity as a species, and 2) that genetic engineering modifies the biology of future humans against their will.⁹³ Implicit to premise 1 is the suggestion that the moral dignity of individuals rests on understanding oneself to be a member of the species. Further, the sincerity and authenticity of action and communication about action are fundamental to Habermas' account of autonomy, which would be undermined by any *distortion* of biology imposed on the agent. The experiments of life, their diverse value as autonomous acts, Habermas seeks to defend against erosion through engineering, and he argues "since we are inexchangeable, we act and judge *in propria persona* - that is our own voice speaking and no other."⁹⁴

In a case where agents could choose A or B, being presented with risks and benefits to their bodily wellbeing, it is not enough to predict accurately what individuals will choose, because

the communication of autonomy involves the sincerity that must be articulated only by the agent. Predictive ability, its use to make beneficial paternalist decisions on behalf of agents, might seem more justifiable when the agent's preference for A or B can be explained solely by reference to external factors, but this still is insufficient. Unpredictable autonomous persons, although the sincerity and authenticity of their acts may be obvious to a lesser extent in some of the activities of life,⁹⁵ are nonetheless those with moral dignity. If an agent should be entirely free to reflect on the same choice and make either of the different and 'inexchangeable', decisions available, then we must reject even the most beneficent paternalism that might seek to intervene.⁹⁶

In one variation of the compatibilist view, according to which autonomy is compatible with influence and should also be reasons-responsive, an agent is autonomous even if they desire A but are prompted to choose B by external forces. If an external factor led to a change in the beliefs and desires that would otherwise have led to A, the agent would have a revised preference for B. When there are good enough reasons for an agent to change their beliefs and desires, an external factor could therefore result in a different action, and an irresistible external factor (a reason) might still be compatible with making an autonomous choice. To continue the example above, the agent could autonomously choose B, when presented with a reason to do so, even if A would have been the more coherent choice given their beliefs and desires, following self-reflection. On this account of autonomy, a change in behaviour, even if the change was motivated by irresistible external forces presenting reasons, does not mean the agent acted under duress. The implications of allowing irresistible external factors to be part of the process contributing to autonomy vary in their significance, and some seem more easily justifiable than others. This variation depends in part on how important it is that beliefs and desires contributing to a particular action should be flexible. In some cases, irresistible external forces might simply lead to individuals acting out of character, incoherently, or in ways they could not easily explain. It might be that in some circumstances it is more normal for preferences to be less susceptible to external influences or contingencies, but this does not mean quickly changed behaviour must be insincere or inauthentic, or that it is not freely changed.

Given an understanding of autonomy which allows the influence of external reasons and direction by irresistible internal forces, it is difficult to identify necessary and sufficient conditions for autonomous behaviour. Even given a capacity for self-reflection on beliefs and desires, practiced in a socially normal way by someone who sincerely and authentically acts on relevant beliefs and desires, there still seem to be grey areas. One difficulty are cases

where the agent appears to act autonomously, given socially normal awareness of their sincere and relevant beliefs and desires, but nonetheless fails to behave appropriately. Consider the socially motivated belief 'daily exercise is good for my health', and the relevant (also socially supported and contextualised) desire 'I want to be healthy'; an agent might sincerely, authentically, and self-reflectively hold these beliefs and desires, supported by self-reflective understanding of the right course of action for oneself, and contextualised by society. This individual might still fail to act, and this is far from an unusual case for autonomous individuals.⁹⁷ Behaviour demonstrating weakness of will seems to characterise some kind of failure of self-control or self-reflection on beliefs or desires, though the capacity for reflection and the sincerity and authenticity of the beliefs and desires remain. Weakness of will, if weak-willed behaviour can be called autonomous, suggests that the matter is more complex than identifying and discounting cases where individuals act consistently or exert self-control based on their beliefs and desires.⁹⁸ Nonetheless, the capacity for self-government seems central to autonomous action, even though self-government might also allow momentarily renouncing self-control. It seems likely that there can be autonomy even when an agent chooses to act against their own best interests, because autonomous agents might lack focused motivation to follow through on certain beliefs or desires in particular instances and might fail to reflect without losing the capacity to reflect.

As well as a capacity for self-reflection, and the ability to put the capacity to use at will to guide preferences and take appropriate action, being able to self-identify as the author of an action seems important for the meaningful exercise of everyday autonomy. The judgement that freedom to act according to one's own beliefs and desires is integral to maintaining one's personal identity might be one motivation for maintaining some concern about paternalism and moral coercion, even having accepted compatibility of autonomy with persuasion, circumstance, and causal contingencies beyond agent control. If an agent's behaviour consistently contradicts their stated beliefs or desires, it might suggest they are either incapable of rational self-reflection, or else indicate that factors other than the agent's beliefs or desires were more relevant causes of their action than their deliberation. In some cases, apparently incoherent behaviour might not be too troublesome for an everyday account of autonomy. In case a decision was likely to be personally transformative, it might be necessary that autonomy should have the requirement of greater sensitivity to reasons and the risks involved. This increased requirement would separate the demands of autonomy in personally transformative cases from autonomy in more trivial cases. It is true that certain behaviour is more usually understood as a reflection of individuality and

freedom of thought than other behaviour. So, the tolerance for incoherence should be correspondingly low in some cases, but this does not mean the actual autonomy at stake changes. Comparatively meaningless or impersonal patterns of behaviour might justifiably be influenced and vary to a greater degree, without causing injustice, even if the influence is an imposition on autonomy. In medicine, the different epistemological status of patient and clinician, and the vulnerability of patients means stricter legislative oversight is important to safeguard against harm, even though the bodily autonomy at stake is more significant and personal than in other cases of autonomous choice. It seems likely in some medical cases that autonomy is compromised slightly by medical care, but that this is justifiable.

For a defence of transhumanism and the nudges that might be required to promote it in the public eye, a focus on views that describe autonomy as compatible with determination by internal causes and some irresistible external influences will be most fruitful. Similarly, given the requirement for informed patient consent in medicine, which is intended to protect a level of patient autonomy, certain views of autonomy will suppose that autonomy depends predominantly on the capacity to reflect on beliefs and desires, with the influence of some external factors acting on that capacity in variously useful and justifiable degrees. In the next section I will aim to show the difference between examples of influenced but autonomous medical choice, and influenced and non-autonomous medical coercion, with a view to broadening this discussion in a later chapter, to consider the justification for some medical nudges.

Is there any hope of a global framework for public healthcare responsibility?

Policy is guided by social values, in medicine as in other national matters. As a result of variation in beliefs and values, as well as differences in wealth, there is variation between the medical standards of best practice of nations. Medical advances, the result of new possible kinds of medical care given emerging technology, or new applications of existing technology, are accepted and incorporated at different speeds in different places. The relatively recent field of global health aims to take the values shared by most and oversee their inclusion in the policy made by all cooperative nations. Healthcare is widely considered a human right and so policies to protect and promote health are increasingly considered a human matter. Because the values of national public health do vary, and because the policy makers in individual nations are guided by different parent (global) organisations, or by none at all, it is hard to discuss public health policy, norms of consent, or the hierarchy of authority

and responsibility on a global scale. For clarity and brevity in the thesis, I will discuss policy making and public health responsibility in the United Kingdom and the United States as two interestingly different examples of democratic societies that are guided by global health organisations in their policy, and who also participate *as* global health authorities to protect the rights of people beyond their borders. For example, the UK Department for International Development (DFID) engages in agenda-setting activities and provides technical and financial support in nations outside the UK, and its counterpart, the US Agency for International Development (USAID) has recently engaged in relief work internationally to support vaccine delivery, citing ‘cooperation’ as one of its core values, and to support women’s rights in Ethiopia.⁹⁹ There are other interesting and useful examples of global health organisations whose work guides policy and protects the health of many, such as the work of Médecins Sans Frontiers, UNICEF, and The Wellcome Trust, whose work I mean to include whenever I discuss the increasing success, prominence, or authority of widely recognised global health organisations in general. The most prominent, because of its breadth and inclusivity, is the World Health Organisation (WHO), who are responsible for widespread agenda-setting, standard-setting, guidance on policy making, and advocacy, with near universal national membership.¹⁰⁰

The organisation of the WHO is such that many stakeholders collaborate, and act cooperatively to carry out global goals and to strengthen the hand of national policy makers. The organisation is not arbiter of global regulations and does not have unilateral authority to enforce standards of health in any nation. However, given its reach and expertise,¹⁰¹ and the established network of local offices and authorities from the WHO globally, the coordination of activities by the WHO is the best example of global health policy, or global standards for healthcare practice, which does not take a narrow or privileged view of healthcare norms, and cannot afford to do so.

Medical care in the United Kingdom is predominantly free at the point of use, and the existence of subsidised care as the *National Health Service (NHS)* has notably been a matter of public pride (rather than simply of public acceptance or agreement, for example) since its inception in 1946. In the United States, health insurance has traditionally been the responsibility of individuals in cooperation with their employer, and public health policies tend to have different authority to restrict and manage the behaviour as individuals. Attitudes to national or socialised healthcare vary but may become less diverse as global health becomes more prominent.

I will establish a chiefly descriptive groundwork to support the arguments that follow in Section Two. The disagreements that follow are intended to present a view of some key concepts in the philosophical literature on autonomous consent, medical paradigm shifts, and emotions as part of reason, to help shape the problem for transhumanism that is addressed in the second half of the thesis.

Whether personal autonomy can stand the influence of external forces or not, the practical application of reasons-responsive autonomy is central to existing frameworks of medical ethics. Given the need for non-expert patients to decide on serious matters of expert care, guidelines for the ethical practice of medicine are guided by the minimum requirement of patient autonomy, exemplified by patients' freely given informed consent.

Informed consent protects bodily autonomy, meaning procedures cannot be carried out without patient permission, but it does not mean the patient is more able to participate in carrying out procedures themselves. The autonomy in medicine is preventative rather than active, to some extent, as the right to say no to bodily interventions, or permissive, allowing others to act. The autonomous ability to give *informed* consent to the actions of specialists also relies heavily on a patient understanding the risks and benefits of available treatments. As a full understanding of the medical risks and benefits involved in treatment plans require years of medical training, the necessary information must be carefully delivered by experts, and its delivery is constrained by legislation to protect patients against undue influence.

Patient autonomy is one of many key constraints for bioethical and medical practice. In this section, the limits of patient autonomy will be explored in terms of its conflict with the other key considerations, here described as the four key considerations of bioethics,¹⁰² and as a practically grounded example of a diverse philosophical literature. Autonomy will be considered as an attempt to describe a conceptual ideal, before reintroducing the limits arising from practice in real cases. In modern medicine, informed consent must be given by patient or their proxy for the autonomy condition to be met and will be considered here with a view to its potential as an enduring guiding consideration for ethical enhancement.

Professional involvement in how individuals make personal medical decisions should aim to protect rather than override patient autonomy. To return to a difficulty raised in the previous section, informed consent in medicine has a lower threshold for persuasion by experts than in other areas of important activity like voting. The aim of professional involvement in securing patient consent is to inform; to present all available treatment options clearly and fairly with the patient's wellbeing and understanding as guiding principles. The consent of

the patient, or their next of kin as consenting proxy, is necessary *before* most medical procedure or investigation can be carried out. The legal requirement is also that consent is given before procedures or investigations, and that this consent should be free and informed. This requirement applies to participation in clinical trials, treatment options during life, some blood and tissue removal and use, and some post-mortem interventions, as well as to matters of privacy, and patients must understand and appreciate what will happen to information about their treatment.¹⁰³ Laws enforcing informed consent are intended to protect the autonomy of living patients, and to protect dignity and privacy after death.¹⁰⁴ The health status of an individual usually cannot be disclosed against their will, even after their death, without positive consent. In some cases, patient privacy can be breached if immediate and direct harm would be caused by withholding information from another patient or patient's close relative, but the use of information is restricted, and such cases are usually controversial. In some cases, the requirement that patients are fully informed about all alternatives or consequences attending a treatment option are less stringent than in others. For example, registration as an organ donor (post-mortem retrieval of tissue) involves less stringent oversight of patient understanding about organ use than do decisions affecting radiotherapy or chemotherapy, or tissue donation during the life of the patient. Though having or acquiring consent (from patient or proxy) remains a necessity after death, the need for discussion of physical risks to the patient from post-mortem procedures are replaced with the consequences for others.¹⁰⁵ Controls to standardise and monitor consent are in place in most countries, as well as communication guidelines for how consent can be legally acquired across medical practices, clinical trials, and for some post-mortem interventions. These recommendations and guidelines are overseen by national health bodies, governments, and by the World Health Organisation, and breaches are treated seriously, even where they are not directly regulated by law. Regulations and recommendations governing how clinicians should communicate with patients to procure informed consent are not always prescriptive and more broadly represent the goals of good practice.¹⁰⁶ Interactions between medical professionals seeking to acquire consent from patients are the subject of careful training and oversight because professionals are not permitted to make decisions on behalf of a patient, even though they provide the information needed for the patient to decide. In some cases, professionals are not permitted even to recommend a particular medical decision over others, although they are the primary source of medical expertise.¹⁰⁷

Given the difficulty of communicating complex information about risk without undue bias, such that it is comprehensible to a non-specialist, some limits on the patient's capacity to

give truly informed consent seem all but inevitable. Nonetheless, some influences are legally seen to undermine informed consent where others do not, and this can be better understood by first returning to definitions of autonomy in the philosophical literature.

Autonomy as a first-person concept is the feeling of control and perceived capacity for self-determination. In medical consent, this corresponds to the patient feeling able to refuse treatment and feeling able to understand the consequences of their choices, to determine what happens to their body. As a second-person concept, autonomy in the context of medical consent means clinicians must respect patients as capable agents with the final say over their own bodies. Clinicians treating capable adults must recognise their patients are able to understand and decide for themselves, despite clear epistemological disparities between clinician and patient, and with sensitivity for the vulnerability of the patient. Autonomy is an imperfect concept in both cases. First-person autonomy is affected by unconscious bias, unreflective preferences, and engrained habits, meaning a feeling of control is not the same as control, and even self-reflection can be fallible. Second-person autonomy is imperfect because patients, as patients, are vulnerable and probably less well-informed than the specialists guiding their decisions. Specialists, knowing this, might be liable to think of their patients as less than fully informed or capable, though the condition of informed consent requires it. These limits on autonomy in practice can be compensated in part by other principles governing patient-clinician interactions; that the clinician should promote beneficence, non-maleficence, and justice. These further conditions for good practice help separate the cases where consent and respect for autonomy are respectably imperfect from those where autonomy falls short, and infringements are made on patient rights and dignity.¹⁰⁸ This distinction could not be usefully described if only the presence of patient autonomy could determine what was good medical practice. Also, although it might be clear that perfect informed consent and patient autonomy are unrealistic in many real cases, this does not remove the legitimacy of informed consent as a goal of good medical practice. Respect for patient autonomy in clinical and medical settings can be pursued only if a sincere attempt to uphold the principle of informed consent is made. However, clinical beneficence and justice are not always straightforwardly able to make up ethical ground where, despite an absence of bad-intentions, autonomy falls short. Sometimes, these other key concerns of the clinician are in active conflict with informed consent or prevent the ideal of informed consent from supporting true autonomy.

In case a patient does not fully understand their options, clinicians have a duty of explanation and clarification. Clinicians should have some understanding of the need for their

explanation to help patients avoid being misled by unspoken framing effects and biases like the status quo bias in the face of hard decisions. The clinician must try to communicate enough information, on the consent form or in conversation, so that the patient can grasp the implications of each of their options even if their understanding of anatomy, healthy lifestyle, self-care, or medicine is limited. Information about a variety of options and risks must be given without privileging any option beyond its expected clinical benefits and risks to the individual. For example, it would not be good practice for a clinician to encourage a patient to choose one procedure over another for reasons of cost, personal preference of the clinician, or to make it easier for the clinic (*i.e.*, because certain equipment was cheaper or easier to maintain). In case it remains clear that the patient does not understand the treatment options or risks after explanation, clinicians might have a duty to simplify their communication. In practice, to avoid discrepancies in communication in general, clinicians or with greater technical understanding than the patient will tend to follow the standard approach of their clinical body (hospital, organisation), to keep patient advice consistent. Consistency across clinicians in presenting the simplified 'minimum' information necessary for patient comprehension, though it does not completely ensure patient understanding of medical procedures, means discrepancies in communication or clinical preference are less likely to direct patient choices. Where full technical understanding of options and risks is too hard, therefore, understanding the personal implications¹⁰⁹ of choosing any option might suffice, if all patients have access to a standardised minimum of information.

Particularly in a case where possible negative side-effects are extensive and complicated, and where some of those side-effects are unlikely but extremely severe, there is a greater risk of presentation effects or bias creeping in. Particularly in cases where patient comprehension of key terms might differ from the contextual understanding available to medical professionals, the result could undermine autonomous choice, as informed consent. Medical good practice might involve more careful *tailoring* of the treatment plan, to suit individual patients (rather than blanket and unilateral imposition of one treatment) and this becomes easier the greater the resources available. Though more personalised and tailored medical treatment and presentation of choices might result in better patient outcomes in terms of care, or feelings of patient satisfaction,¹¹⁰ it also presents more opportunity for bias and inequity.

When a patient must choose between treatment options, the patient should expect to need to pay attention to the clinician and accept their assessment of risks and benefits of the available options. It is normal for the clinician to present information about even moderately

uncommon risks carried by treatment, as deliberately overlooking side-effects of treatment would seem not to allow the patient truly informed consideration of their choice

If one of the very unlikely risks of treatment necessitates the use of very loaded terms (*e.g.*, cancer, herpes, morbid, radiation), the patient might respond more strongly to loaded language, or to perceive as less risky options with obscure or neutral-seeming terminology. Where patients misapprehend risk based on familiarity or apparent seriousness of certain words, their information about the risk seems to fall short, but not in the same way as where a risk is deliberately undisclosed. Misapprehension of high risk might result in patients choosing against their best interests.

If patients avoid treatment because of the emotive language used in explaining the undesirable outcome, this does not necessarily undermine their competence to decide for themselves, or to give consent, because their emotional responses are their own. Competence to decide might be understood to result from both understanding and appreciation – understanding of the facts and appreciation of the significance of choice.¹¹¹ In this case, the effect on competence of a decision based on a response to loaded language would need to be further examined. If the effect of loaded language on the reasons for a decision is the result of a complete misunderstanding of the information presented, this would undermine competence to decide. If the effect of loaded language on the reasons for a decision is the result of a misapprehension of the relevance of the facts presented to the patient, this also seems to undermine competence to decide. However, Patient B does understand the facts presented, and appreciates the relevance of each risk to their case, if they have paid attention to the clinician. It is not clear that their competence to decide is eroded, if they understand and appreciate, even if they also have an emotional reaction to particular facts or language. The way information is presented to patients is not sufficient for informed consent, therefore, even if the intention has been to minimise the risk of bias.

In a 1960s Chronic Disease Hospital case,¹¹² clinicians removed loaded language (mention of ‘live cancer cells’) from the information given to patients. In this case, the aims of the clinical test were also miscommunicated deliberately, to offset an expected bias against certain ideas. Patients involved in the trial were already vulnerable, and the double miscommunication of information about the reasons and method for the procedure meant they were not able to understand or appreciate what was done to their bodies.

To offset the effect without undermining patient competence, a clinician might continue to use the loaded language, rather than removing it from the patient information, but attempt

to offset its perceived riskiness in other ways. Framing effects are understood as a common bias affecting risk assessment.¹¹³ Patients presented with a positively framed option are more likely to be risk-averse, where those presented with a negatively framed option are more likely to be risk-seeking.

In Tversky and Kahneman's classic study of framing effects, participants were asked to imagine that *a disease was predicted to kill 600 people. If Programme A were adopted, 200 people would be saved, but if Programme B were adopted, there would be a 1/3 probability that 600 people would be saved and 2/3 probability that no people would be saved.*

A second group of participants were given the same choice, framed differently. *If Programme C were adopted, 400 people would die. However, if Programme D were adopted, there would be a 1/3 probability that no people would die and a 2/3 possibility that 600 people would die.*

The positive framing in the first pair of choices, '200 would be *saved*' appeals more often than the same result presented in the second pair, '400 would *die*'. This framing effect has been found to make Programme A more appealing than Programme B, despite the same numerical loss of life.¹¹⁴ Knowing this, clinicians attempting to offset the loaded language in one option by use of deliberate positive framing would not be interfering with competence. Such minimisation of unlikely but alarming consequences seems a more realistic and appropriate as part of a clinician guiding patients than withholding information, because it does not undermine understanding and appreciation.

Given the well-documented effects of framing and status quo biases, anchoring and other effects on behaviour, an impartial patient response might seem very unlikely. Presenting risks as fairly and transparently as possible, to secure consent for procedures they judge beneficial, is a duty of the just clinician. Carrying out this duty responsibly requires some awareness of common bias and presentation effects in the way minimum information and simple communication strategies for medical practice are standardised and regulated. As clinicians are fallible too, truly neutral presentation is unlikely, and attempting to navigate serious personal decisions with flat affect does not seem likely to improve either patient or clinician autonomy.¹¹⁵ Given these known communication difficulties, if clinicians anticipate common patient reactions to certain procedures or language and aim to offset apparently irrational effects, this seems to uphold justice. Offsetting bias in doctor-patient interactions means an intention to improve parity between rational agents with autonomy to choose but does not necessarily mean giving the patient any more or less information than seems

conversationally appropriate to the doctor at the time, because clinical interactions are private and because the best judgement of the clinician is trusted.

In patient-physician interactions, the patient has *'the right to be involved in decisions about their treatment and care and be supported to make informed decisions if they are able.'*¹¹⁶ In communicating information to patients, to allow their involvement in a decision-making process that often involves a level of medical expertise unavailable to the layperson, it is the duty of the physician *'to try to find out what matters to patients so they can share relevant information about the benefits and harms of proposed options and reasonable alternatives, including the option to take no action.'*¹¹⁷ The assumption made by the Guidance is that the expertise of the physician must not override but rather must consider the values held by the patient when laying out options for the patient to consider.¹¹⁸

The GMC Guidance suggests that the information physicians should provide to patients who must decide what to do, will usually include the following:

- a. *diagnosis and prognosis*
- b. *uncertainties about the diagnosis or prognosis, including options for further investigation*
- c. *options for treating or managing the condition, including the option to take no action*
- d. *the nature of each option, what would be involved, and the desired outcome*
- e. *the potential benefits, risks of harm, uncertainties about and likelihood of success for each option, including the option to take no action.*¹¹⁹

This Guidance helps identify a challenge involved in applying the seven principles that are well defined and accepted in bioethics for general medicine to the behaviour of policymakers in public health settings. Individuals have the right to be involved where diagnosis and prognosis suggest that action must be taken, with different consequences, and where there are potential risks of harm and uncertainties about the likelihood of success, which would change the preferences of the individual, even if the ultimate and unchanging goal of individual health and wellbeing remains in place for both patient and physician.

When the bias effect is toward loaded language, an additional question might be how to offset bias in a comparable way to positive/negative framing, without changing content. In a conversational context, other cues might be available to clinicians to approach sensitive words without omitting or changing content, and it is easy enough to imagine that this would

constitute something like a clinician's duty to compassionate treatment (their bedside manner), which is also generally understood to be beneficial to patients.¹²⁰

Separate to the difficulties faced by clinicians who need to provide face to face communication, consider the comparative simplicity of the information communicated via standard written consent forms. WHO consent forms given to participants ahead of their participation in clinical studies aim to give a clear explanation of the reason for the study, some relevant elements of its design, and for the method of participant selection. Clarity in these areas is in part intended to assuage fears about why particular participants were selected to take part.¹²¹ Some details of methodology and design might mean participants learn they might not receive the test drug (*i.e.*, that some participants will be given a placebo or an alternative treatment, and that this will happen at random, perhaps under double-blind conditions). The consent form does not attempt to offset the risk that participants are likely to feel they are more likely to be given the test drug than the placebo. This is a known bias, with participants in general more likely to believe they will be luckier than average, despite factual understanding that placebos are also being randomly assigned to participants with identical odds.¹²² In written communication to procure informed consent information must still be presented clearly and without emotive or persuasive language. Unlike a face-to-face conversation, when information is primarily given in writing, bias and emotional responses cannot be addressed without changing content across all participants.

Finally, significant limitation of patient autonomy can happen beyond presentation effects, problems with understanding, and emotional effects, because of the patient's situation. Expectations about health, standards of medical care, and patient freedom to refuse treatment or disagree with clinicians are affected by wealth and social status in many cases, even in countries where healthcare is free at the point of use

Even in systems with more comprehensive social and medical care nets in place, more indirect financial pressures continue to affect patients and clinicians. Where certain procedures are limited by location, by the number of specialists, long waiting lists for particular procedures, or other structural pressures, even if healthcare is free for all at the point of use, demands on the health service provider mean restrictions.¹²³ The principle of justice could protect patient autonomy from restriction of this kind. For example, transparency about the associated risks of joining a waiting list that is too long given the urgency of a patient's situation seems like information that should be part of the consulting process, though this is a social and political rather than strictly medical factor. Informed

consent, given on the basis of a procedure being available immediately is different from consenting to be placed on a waiting list for an indeterminate time, during which the patient's condition could change or deteriorate. Information provided should concern assessment of *true* risk of the procedure to the particular individual, including social and political obstacles to health.

In clinical settings, when a medical decision is presented to a patient, the need for the decision and its risks and implications should be clear, and clinical priorities should be made explicit. In everyday action, self-transformative and momentous instances of autonomous choice can be less obvious, and are not necessarily 'presented' at all, let alone presented in terms of their relevance to particular social values. Public health matters like healthy eating and lifestyle changes also tend to involve the latter, less overt, choice architecture rather than transparent and carefully managed clinical decisions. Common examples like snack choices, financial planning, and stair use, as demonstrated in typical nudge cases, can be manipulated in ways that do not rely on clear communication with individuals, in a way that would not be acceptable in a clinical setting. Although public health matters do share some values with the values of medicine, it is not clear that the demands of communication and autonomy in the clinical setting should also be required in the public setting.

Questions for transhumanism

Democratic libertarian societies prioritise the informed free action of individuals in matters concerning their own bodies, where action does not intervene on the bodily autonomy of, or cause unjust harm to, others. However, a libertarian society can nonetheless have governing health bodies, detailed laws governing appropriate clinical and medical procedures, defining malpractice, and train medical professionals who are entitled to act on their beneficent judgement for patient wellbeing, to promote justice, in some cases.

One tension for philosophers interested in influencing public health is that certain accounts of personal autonomy allow more intervention and more robust use of government and social pressure to influence personal decisions than others. True autonomy might appear to be free from influence, but in reality, socialisation makes this unlikely and impractical. Key questions for transhumanism arise given an account of autonomy as informed consent, as practiced for medical ethics, which seems to be an autonomy open to intervention from external pressures by necessity of the situation. In medicine, patient autonomy must be left

intact in the form of informed patient consent with proper engagement at the appropriate time so that individuals can apprehend the significance of their choices, the interactions of their choices with those of other people, and the effect of their choices on others (particularly in cases where genetic information and heritable illness is concerned).

It is important that patients are consulted in clinical settings at the right time, in the right way, with careful management of inevitable clinical pressure and social pressure to allow patients to avoid coercion and choose well. Although I will suggest that the spirit of inclusion, dialogue, and management of clinical discussions forms the foundation for ideal public health engagement, the regulation around seeking informed consent, which depends on the input of beneficent experts, though reasonable in medical and clinical settings, would be very demanding in other settings like public health. As transhumanist interventions on the body will tend to involve personal medical decisions taken by individuals, but the influences guiding personal choices are inherently social and political public health matters, then the following questions arise for transhumanism in clinical settings and as the subject for public debate:

1. Is it necessary for non-experts to give informed consent to participation in a project of species-wide enhancement? Given the likelihood of as-yet-unknown future implications of enhancement, is it even possible for the layperson to consent?
2. Can patients choose transhuman enhancement or receive expert guidance in the same way patient choices receive guidance in general medical settings? To what extent do behavioural insights and nudges offer valuable support to general medicine in the public health realm, and could this act as a template for the approach needed by transhumanism?

It seems coercion is a threat to autonomy, and the concern shown by the WHO for patients and participants in clinical trials, who are already vulnerable during their interactions with clinicians who must inform, means they approach consent carefully. However, critics of nudges argue¹²⁴ that autonomy can be similarly vulnerable in a social and political setting where the liberties of free thought and action are guided by norms and biases that might pass unnoticed, given their familiarity. The vulnerabilities of patients and citizens also overlap, given norms of collective responsibility to contribute to public health, which are

influenced by political decisions and social pressures. Collective responsibility to others, and the unconscious effect of implicit and cultural cognitive habits limits and guides autonomy and often frames what can be considered respectable decisions. These influences, conscious and unconscious do constitute a kind of beneficent external pressure to behave a certain way, about which Mill was mindful. The need to move beyond custom, and to be guided by a shared responsibility to do well by others is expressed in the harm principle, that “the only purpose for which power can be rightfully exercised over any member of a civilised community, against his will, is to prevent harm to others.”¹²⁵

However, as becomes clear in medical examples of good and bad practice, particular kinds of external pressure on the autonomous individual do not mean awareness of shared social expectations, or reminders about group responsibility always damage individual autonomy. In terms of both promoting and implementing species-wide enhancements, it is from political and social influence that the worry about an unjust external pressure on human beings to become homogenous in body, thought, or action, or to live in bodies augmented by the will of others, becomes most challenging.¹²⁶ In chapter 6 these pressures will be considered only in terms of their impact on autonomous choice, applying the established behavioural insights paradigm to practical transhumanist cases. Broader concerns about how promoting transhumanist ideals might create unjust pressure on people to become homogenous will be addressed in chapter 5.

In the second dialogue of this chapter, I will discuss the ethical implications of changes in medical knowledge, given the influence of technology on policy and social norms. In the existing paradigm, there is a distinction between treatment and enhancement, which is pertinent to understanding transhumanism as a new public health paradigm with clinical implications for the individual. Consensus is absent, about how the distinction between treatment and enhancement should be drawn, which is both an influence on norms and policy and is influenced by norms and policy about existing technologies.

Second Dialogue: On How Medical Knowledge Should Change

“Authority-based medicine, which has always been a foundation for medical knowledge, is ancestor of the medical consensus conference, which also aims to produce knowledge based on trust of authorities. The difference is that authority-based medicine rested on single

authorities (say, Hippocrates or Galen), while consensus conference-based medicine makes use of the rhetorical power of the consensus of a group of experts.”¹²⁷

Consensus conferences do not always reach a consensus. When this is the case, perhaps because the matter under discussion is controversial, or where there is ‘insufficient reason to diverge from current practices’ even if there is evidence that another practice might be an objective improvement, the results are still reported to the press, in a spirit of openness. Solomon notes, however, that consensus conferences must try to balance interesting controversy with uncontroversial ideas, openness of discussion with some privacy for executive discussion, specificity of outcomes but not to the extent of prescribing processes to medical practitioners entirely, and while taking time to consider issues in depth but without running overlong (for either expert participants or for the press who must hear and report on results to the community).¹²⁸

Critics of consensus conferences suggest that their methods are unlikely to produce unbiased results, given pressures of time and publicity, and because participants tend to be experts whose views tend toward the conventional opinion on matters under review. The consensus conference approach of weighing evidence from scientific trials against the predicted outcomes of changes in policy and practice may be practically realistic but, critics suggest, is not a rigorous way to arrive at truths, nor to disseminate best understanding to the press (and so to the public).¹²⁹

Though consensus conferences may be inefficient, unlikely to arrive at rigorous evidence-based objective results, they do highlight a need that exists even among experts to discuss reasons and the value of likely outcomes, and to engage in rhetoric about best practice and convention while so doing, rather than instigating policy change from quantitative results alone.¹³⁰ Evidence alone does not tend to form the basis for policy decisions in part because it requires the opinion of experts who are familiar with the existing landscape of practice and policy. The problem with the opinion of experts in convention, even in the context of deliberating over what to do based on new quantitative results, is that opinion is bias-prone. Supporters of consensus conference suggest that, though the events may not produce pure unbiased results directly from meta-analyses of data, the rhetorical component is obvious, and disagreements about interpreting best practice based on practical considerations from the status quo can be more obvious. The interface consensus, where expert opinion meets society, is also important though it does not seem to occur at the right moment for academic consensus to be at stake in the discussion.¹³¹ The accountability and openness of a publicly

presented expert opinion resulting from the consensus, as though from science to society, also seems valuable for the layperson who would not have time to consider evidence, and has limited freedom to avoid the consequences of such consensus as it affects practice.

Different approaches are appropriate during earlier stages of clinical trials than later, where later clinical trials rely more heavily on meta-analyses from evidence. 'Medical institutions can be designed to minimize potential conflict between the methods (for example, by holding consensus conferences after formal evidence review).'¹³² Though the academic consensus may already exist at the point of the discussion, the report from the conference to the media has its place in shaping narratives that may be meaningful in patient-clinical interactions and can shape public debate.¹³³

Democratic debate on ethical issues resulting from medical science seems to rely on the public being able to engage with the implications of medical knowledge, at least the extent that policy decisions can be contextualised, the use of authority or legislation legitimised by the spirit of an open discussion about why measures are necessary. Further, the results of consensus conferences, though they may not avoid individual biases (Solomon gives the example, whether the panel of experts slept well the night before discussion) they can at least aim to avoid the bias of commercialisation of results, or government pressure to produce results that align with policy.¹³⁴

In the introduction, I presented as an assumption of the thesis the idea that transhumanism is now inevitable. In genetic modification, genome editing, age-reversing biotechnology, prosthetics, artificial intelligence, brain-computer interfacing, and in medical cognitive enhancement, the recent rate of progressive change suggests transhumanist goals are more closely aligned with social goals than ever. However, inactivity in medical ethics is not an adequate response to a distribution problem, given the possibility of making interventions to reduce harm. Inactivity would not reduce the number of possible problems resulting from enhancement but only mean that the kinds of problems experienced were not predicted, controlled, or the obvious responsibility of anyone in particular.

Inactivity may reduce the risk of guilt for those who would aim to make a difference and could fail, but it seems unlikely to result in an overall better outcome than deliberate forward planning. Existing defences of intervention have suggested that the higher cost of faster success is justifiable when the investment means more lives are saved through faster implementation of lifesaving innovations.¹³⁵ In practice it is hard to determine with accuracy the cost of delaying lifesaving research, even against the acceptable number of global deaths

from a given cause. Particularly when the goal of the research is to change the average life expectancy and normal range of health outcomes in a target population, measuring acceptable success or determining what would count as failure poses a problem. Because the success of transhumanist enhancements should mean the average life expectancy (perhaps also cognitive capacity) of the population increases, successful research and development might at first create an innovation that means more lives end before they must.

For example, early dialysis machines were costly, large, and could only be operated in hospitals by trained medical staff. Patients were, at first, chosen selectively by a committee of citizens who assessed the suitability of individuals to receive a chance at life-extending treatment. The existence of the machinery presented a choosing problem, and the initial committee process would later be revised, and considered hugely flawed, in part because the deliberations were carried out without transparency and on the basis of value judgements relating to social and economic class. The committee process was advantageous over alternatives such as ignoring the potential of the limited number of dialysis machines to treat a limited number of patients, because of the unfairness of distribution, or else ignoring the need to select patients carefully and selecting randomly on the grounds that choice would be bias-prone.

Instead of measuring the success of an enhancement intervention in terms of lives that would be lost or saved from premature death, where the definition of premature or unnecessary death is changed by the same technological innovation, the high cost of carrying out research into new medical technology should be very carefully balanced against the need to fund existing medical therapies but should be considered a part of the process. It would be hard for medical research bodies to justify redirecting funding from established and successful cancer treatments entirely into funding for genetic engineering, even if the enhancement would reduce cancer rates overall because the immediate demand for treatment cannot be paused, and because current patient outcomes are time sensitive.

In part because of the predictable short-term loss of life that would result from reducing the support for established therapies like cancer care in order to fund research into novel treatments, such a redistribution would be understandably unpopular and would cause unnecessary suffering. Though in practice it is highly unlikely that resources would be taken directly from therapeutic medicine for enhancement in this way, the need to distribute resources appropriately remains, and expediting transhumanist enhancements may not be justifiable in these terms. The deliberation of experts for example, might result in a practical

consensus that areas of medical research were promising, but that the costs to existing care could not justify a change in priority. In jurisdictions where there are already concerns about the nationalised financial support for the practice of general medicine, and where practitioners do not necessarily have access to the most up to date training in conventional medicine, it seems more likely that any plan to direct resources towards human enhancement should be rejected on utilitarian grounds.

Because of these concerns, and the consequences that enhancement is more likely to become available in richer nations and only to privileged citizens first, it is important to ask whether and when it is ethically acceptable for governing bodies to intervene to expedite transhumanism. Reports on the ethical status of heritable gene editing have begun to suggest that it is reasonable for existing medical practice to introduce more interventions like genetic screening as a part of routine early life care because this would reduce disease burden relatively quickly. I suggest that ethical egalitarian distribution relies on the action of governing bodies, to address any public perception of enhancement as separate from treatment or less valuable.

On paradigm shifts more generally

When enough of a discipline's initial assumptions and practices no longer match the prevailing understanding of the world, it is time for a change. This seems to be true in practice, even if it is not possible to talk about achieving accuracy in terms of true explanation of natural phenomena.¹³⁶ Certainly, in terms of medical science, the prevailing idea of health is important and discrepancies between medical science and general accounts of health cause friction.

Kuhn's account of change in the sciences offers an explanation of where paradigm revolutions should happen, and why they would mean progress, as well as why a new exemplar is needed if the best available explanation is to make sense of the world.

In the study of the natural world, evidence is gathered, theories adapted, and new assumptions replace old, furthering scientific progress. Having introduced Kuhnian paradigm shifts, which are still the basis for accounts of change in science, in the second piece of the dialogue I will explain why recent criticisms of the way change comes about in terms of paradigm shifts can help us make better sense of modern medical science and ethics. In the third part of the dialogue, I will give an account of paradigm shift as it is likely to affect an

understanding of ethical transhumanism and raise questions for transhumanism to answer about good ethical practice.

The Kuhnian account established that periods of different activity make up science, and that there should be different expectations of method and practice in each. Scientific understanding during the period of 'normal science'¹³⁷ generally should be consistent with established assumptions and new evidence should be incorporated to refine normal scientific practice. The extent to which individual people are expected to accept new evidence varies with their level of expertise, as experts usually change their beliefs long before the effects of progress reach the layperson. Within normal science it is usually understood that new evidence should change scientific understanding, but this does not mean established assumptions should be forever in doubt in conditions of normal science. Doubt brings an end to normal science, when the model of established assumptions is thrown into doubt, bringing on model crisis and model revolution. Eventually, accumulated changes in understanding mean the underlying assumptions of normal science should change, in the final phase which Kuhn calls paradigm shift. Before the paradigm shift, normal science still improves as technology and measurement improve, and this tends to be followed by conceptual changes, which finally mean assumptions are thrown into doubt. It is the changes to theory and assumptions which are most likely to affect what is understood by non-experts, and how normal science reaches the general population. A paradigm shift is required when new facts and new concepts mean common understanding no longer matches the best available explanation of the world. Though this is a very summary account of how normal science can improve over time, it offers an ideal for progressive change in the natural sciences, where evidence-gathering allows better understanding of particular mechanisms, which in turn means better theoretical approaches, and then more effective evidence-gathering.

Before Kuhn's account of the incommensurability of scientific paradigms, one generally accepted model of cumulative progress explained that progress was the result of new discoveries *adding* to previous knowledge. Aiming to sharpen assumptions and arrive at a truer approximation, the ideal for science was that it should operate with openness about core assumptions, even within a period of 'normal science' and so build on knowledge. The critical reaction to Kuhn's account of paradigm shift, and incommensurability of the old with the new paradigm has been concern about the implications for scientific method. In short, the concern is that science behaves somewhat unscientifically if it operates without doubting its assumptions. Greater awareness and more continuous doubt about the base of

assumptions underlying practice might seem to be a better ideal for scientific enquiry and more likely to promote progress¹³⁸ but this seems much too demanding. Far from suggesting that the ideal science involves blind adherence to the assumptions underlying practice, Kuhnian paradigms instead give a more realistic account of how normal science is possible. The cognitive capacities of the scientist rely on certain unconscious and intuitive skill development, rather than on adherence to a set of formal rules. These quasi-intuitive cognitive capacities are acquired through practice and so are developed over time and field specific.¹³⁹ Reliance on quasi-intuitive cognitive capacities like reference to an exemplar or use of analogy to interpret the problems within a field, rather than being a barrier to openness and rigour in normal science, allows it to take place.

Transhumanism as paradigm

Appropriate and timely model revolution and paradigm shift is not only important for accuracy of measurement or better acquisition of facts about the external world but is prerequisite to perceiving the world. Understanding the world, and conducting normal science according to a paradigm, means not only can we acquire information but are able to make sense of it. With this kind of perception of the world comes the ability to make predictions and judgements, but there is also “a sense in which, even in science, we see what we expect to see.”¹⁴⁰ Kuhn identified that normal science can be prone to self-confirmation bias, where predictions and judgements will tend to bring about discoveries supporting, rather than undermining the paradigm. Transhumanism suggests that human beings need not age and die, and that core capacities need not remain fixed.¹⁴¹ In general medicine, therapy relies on core assumptions that are not obviously compatible with these transhumanist arguments about human wellbeing, though medicine also aims at human health, longevity, and wellbeing, just as transhumanism does.

Kuhn’s incommensurability, in which different views of the same mechanisms and phenomena can be differently described given ‘a transformation of vision’¹⁴² helps make sense of the overlapping values but different conclusions drawn, between medicine and transhumanism. Innovations and changes that make up practical transhumanism often seem to be improvements in medicine, and transhumanism might be described as having a large stake in general medical progress despite different assumptions. The driving ‘vision’, transformed because it relies on reference to different exemplars and analogies, is central to understanding paradigmatic progress. Projects aiming to find out about the world are not

haphazard general fact-finding missions building on overall approximations of truth, but are ordered disciplines, like molecular biology, astronomy, and quantum physics, which come with their own terminology and methods for dealing with the same natural world. It might be tempting to describe transhumanism as a concurrent discipline to medicine, because it sees and deals with the same illnesses, human bodies, and physical limitations, though approaching them with different aims. However, it is more accurate to describe transhumanist efforts as the start of a paradigm shift in medicine. Transhumanism as a paradigm shift in medicine involves a transformed vision of the human body and certain core human capacities,¹⁴³ and is not a cumulative point in the separate field of general lifestyle-enhancing technologies (encompassing almost everything from fire and the printing press to pourable concrete and the personal computer).

Modern biological human bodies at birth would likely be instantly recognisable as human to our very early ancestors. We remain relatively unchanged in this respect in a way that is not mirrored in society, in clothing, technology, norms tastes or preferences, or even the appearance of our homes.¹⁴⁴ This might indicate that such a biologically transformative medical paradigm shift has not occurred for some time, or perhaps never on the scale proposed by transhumanism, and might indicate why the general reaction to transhumanist change might be one of trepidation. However, in the face of increasing evidence that it will soon be possible to make radical improvement to human core capacities, a refusal to acknowledge the need for corresponding changes to concepts of health and illness across medical science seems short-sighted. In Kuhn's terms, the old paradigm which accepted gradual ageing resulting in a shorter health-span no longer accurately represents how the world seems to be, and so is bound to break down to be replaced by a new paradigm. The increasingly well supported transhumanism paradigm, of healthcare and wellbeing, assumes instead that human bodies can be repaired, and that human ageing is not inevitable. The broad paradigm, containing assumptions about death and damage, would need to rest on a narrower 'exemplar' to guide practice. Particularly in the period of model doubt and model crisis, identifying exemplars and good analogies to inform skill development and intuitive practice is more difficult, by Kuhn's account. A particular achievement, demonstrating the truth of a core assumption, which would drive progress and motivate the pursuit of the new (broad) paradigm, might mean the beginning of the paradigm shift, and the end of crisis. In the case of transhumanism, for example, it seems likely that an exemplar could be the first instance of a successful age-repair trial in human beings, to replicate the success seen in some rodent trials.¹⁴⁵ It also seems likely that the transhumanist movement is an instance of

model doubt, approaching model crisis, but that it is also a paradigm which encourages faster movement through the cycle of doubt, crisis and revolution as well.¹⁴⁶

Historically, the development of new technology has tended to prompt concern about escalation of risks. Whether about the risks of conflict, following the atom bomb; about loss of social cohesion or the nuclear family, following the birth control pill; or fear of stagnation of social values in favour of superficial improvements in the lifestyle of the already comfortable, following the increasing availability of personal computers since the 1970s. Technological improvements of any kind tend to raise questions about whether human nature means we will simply repeat mistakes but with more firepower or greater mining capacity, and that some underlying tendencies will mean no paradigmatic human change should be expected. Having identified that this concern might be driven by perception of a lack, that *"It is not enough to 'use technology with a deeper understanding of human issues', or to 'dedicate technology to man's spiritual needs', or to 'encourage technologists to look at human problems'"*, BF Skinner proposed a solution: *"What we need is a technology of behaviour."*¹⁴⁷ The particular weak-point in human understanding identified by Skinner is one he calls the result of retaining concepts of freedom and free will in our understanding of our own behaviour, for evaluating punishment, responsibility and good behaviour. The stagnation comes from holding on to comfortable ideas of why we act, long past the point where neuroscience appears to support their existence, he suggests. This claim about retaining an idea of free will remains particularly controversial, because it seems so reductive.¹⁴⁸ However, Skinner's general concern about how recognisable our current understanding of human nature would be to ancient philosophers might remain compelling to the transhumanist interested in changing deeper human understanding as well as physical attributes of individuals.¹⁴⁹ An apparent stagnation in human self-understanding would certainly make sense of apparently sluggish global responses to existential risks, and deeply entrenched reluctance to respond appropriately as a species. This concern was also at the heart of Enlightenment philosophies, in the form of corrective leaps to modernise some commonly held beliefs about human nature at the time. The shape of the problem, then, is that there is less motivation to adjust parts of our understanding of the world than others, and that this rests in part on comfortable and familiar feelings about human nature that should long since have been considered out of date.

Kuhn identified that normal science 'proceeds on the basis of perceived similarity to exemplars', and for medicine the exemplar of long *and* healthy life has become the general standard for Europe and the West. Transhumanism conceives of the same puzzle-solution

situation facing medicine but conceives of it differently. For Kuhn, assessment of scientific theory involves comparing the theory against the paradigm, instead of applying rules and method to the theory.¹⁵⁰ Scientifically minded ideologies can be seen to rely on 'relations to the paradigm' as guides, and this model of assessing vision and progress, in terms of shifts that could be likened to a Gestalt switch, seems apt for an assessment of transhumanist concepts of i.e. health, ageing, wellbeing. This allows for recognising problems with the current transhumanist approach, and room to allow that its idealised modes of being are imperfect, perhaps even prone to some of the same failures and subjectivity previous approaches. Insofar as transhumanism does involve interpretation of human nature from a particular time and place, it is bound to be coloured by myth, history, and belief as well as by empirical consideration of the raw problem. An idea of transhumanism as paradigmatic, as a way of thinking about certain theories underlying medical progress, accepts and builds on this weakness of perspectival subjectivity, and resists more brittle claims about what is absolutely the best for all human beings.

Though transhumanism relies on making predictions and guiding efforts toward the next improvement in the human condition, its approach is paradigmatically a more flexible one, which values change. The Transhuman Declaration itself, first drafted by Max More et al. in 1998, has been edited many times since its conception, and should be expected to change again. Although in almost every instance the phrasing tends to avoid specificity about individual technologies of bodily change, changeable current problems, or current causes of disease burden, there are changes in nuance. Higher priority enhancements, and the reasons for wanting change vary. The most specific parts of the declaration appear in item one, which mentions the transhumanist stance on ageing, interplanetary travel, and artificial intelligence:

(1) Humanity will be radically changed by technology in the future. We foresee the feasibility of redesigning the human condition, including such parameters as the inevitability of ageing, limitations on human and artificial intellects, unchosen psychology, suffering, and our confinement to the planet earth. (From July 1998 version 2.4)

And the same item, from the slightly earlier unpublished version, 'Transhumanist Principles' which also mentions inter-planetary travel and habitation:

(1) Humanity will be radically changed by technology in the future. We foresee the feasibility of redesigning the human condition, including such parameters as the inevitability of ageing, limitations on human and artificial intellects, unchosen psychology, suffering, and our confinement to the planet earth. (From March 1998 version 2.1)

Inter-planetary habitation and travel, still described as ‘overcoming our confinement to planet Earth’ remain in the contemporary 2002 and 2009 versions of the document, along with continual mention of overcoming the inevitability of ageing. New to the more recent versions are the mention of ‘existential risks’ in place of the threat of technological wars, and of ‘autonomy’ in the place of control over one’s own life. Although the most recent online version of the document is now over a decade old, its restraint from mentioning particular current technologies means it remains perfectly applicable. As regards the proposal that ageing will eventually be removed from normal human life, for example, the document would become out of date, and this kind of revision might mark the end of the usefulness of the transhumanist paradigm as it stands.

One other reason for thinking of transhumanism in terms of a paradigm is that it hints at how, although development and evidence behind theory are the realm of experts and specialists, nonetheless success relies on the results of any shift being accessible to the public in some recognisable form. The form of the paradigm in the public *mind’s eye* need not be accurate or detailed in every respect but should capture the spirit and implications of the new vision intelligibly. With transhumanism, the effects of gradual proliferation of a new way of thinking about human life can be seen in the integration of ideas like a species-wide success or failure, only conceivable given the advent of a fully connected online globalism and the increasingly prominent idea of existential risk, which trickles down into public consciousness.

Another effect of transhumanist approaches to human bodies can be seen in do-it-yourself biohacking and of genetic information being acquired and stored as a commodity, as the result of wide genetic testing for fun and curiosity instead of out of medical necessity¹⁵¹. The former is generally shunned as unnecessary, dangerous, and irresponsible, while the latter evokes a more complicated public response. The apprehension of and acceptance of particular technologies and modes of being, although they are not mentioned by name in

the Transhumanist Declaration, form a large part of the way we can best understand a gradual paradigm shift. If transhumanism offers new ways for individuals to think about their responsibilities in public health concerns and gives new power to individuals to control their own bodies and lifestyles than in generations before, then its influence is likely to be more ethical in general. That the public come to understand the impact of certain technologies and their implications for future human wellbeing is important for the implementation of, still niche, technologies to become more widely normal. Accepting the new paradigm, for the layperson, means referring to new exemplars for human life which are not necessarily found in the values of general medicine. Although these exemplars could also be described in terms of beneficence, justice, cure of disease, return to health, longevity, their apprehension as transhumanist is the Gestalt switch that changes the phenomenal world, though gradually.

What are the parameters for ethical transhumanist success?

As the thesis asks whether governments or governing health bodies can be justified in intervening to expedite the process, leading to successfully realised transhumanism, it is also important to be clear about what such an intervention would involve. Policy work carried out by the Behavioural Insights Team (BIT)¹⁵² produced several frameworks for making large-scale changes to public behaviour with minimal intervention. Given conclusions drawn from the autonomy dialogue that the best approach is one where minimal paternalist interventions are made, even if public health is at stake, an approach that needs only minimal action to bring about large changes is attractive. If an option is Easy, Attractive, Social, and Timely (EAST), it is generally more likely to happen than an option that is more complex, unattractive, that goes against social norms, and seems alien to current values. The behavioural insights work carried out by the BIT aimed to make certain behaviour (*e.g.*, donate organs, quit smoking, eat healthily, invest in a pension scheme) more easy, attractive, social, and timely, and were successful.

However, this does not mean that all social progress can be explained in terms of what is easiest, most attractive, most socially supported, and most timely. Medical and scientific advances sometimes challenge values and norms, and significant social change can happen against the tide of public opinion: revolution is hardly the easiest way to behave but clearly yields results.

An explanation for variations in the rate of social change comes from Cass Sunstein. Sunstein identifies a behavioural scale, on which individuals operate comfortably with varying degrees of norm or peer support. “The zeros go first...”, Sunstein explains.¹⁵³ In this example, Sunstein’s ‘zeros’ who act first are those people who need little or no encouragement from society or their peers to take action they desire or value. This willingness to act does not guarantee their independence will produce particularly constructive kinds of behaviour, only that they will be quickest to act without the support of others. Sometimes, the action taken by the zero will be progressive and constructive, for example in challenging misuse of authority or fomenting revolution about injustice, but there is no deeper reason to suggest that this should always be the case. *Ones* begin to follow enough zeros given the right conditions: once there is enough action to suggest they would have support, people who need a little backup from others will act on ideas they reflectively do believe to be right or valuable. Sunstein’s *twos* might follow enough other people acting around them, but some will never act. Sunstein explains that those people who will never follow the cues of other people acting around them might be stayed by loyalty to tradition, by fear, or by motivation to keep the status quo in place (for example personal benefit). An individual’s place on the scale between first-movers and never-movers is not necessarily a reflection of particular values, though quick movement is more likely to correlate with being comfortable with progressive action than conservatism. Various internal biases, including a general preference for the status quo, regardless of whether it is the best or most rational choice, are well documented motivations for behaviour. Most people can very often be relied upon to act according to the default, and Sunstein suggests that comparatively few people are typically first-movers.

This raises another set of questions for transhumanism:

3. What are the social norms and expectations of health that make bodily enhancement and longevity unappealing, despite improvements in medical science?
4. Are there good reasons for maintaining any of these norms and expectations about health, even if they made the transition to a better public health paradigm less likely?

Another general explanation for slow progress in human self-understanding can be found in a common difficulty for medical ethics; if there is individual responsibility for public health,

vaccination must be a moral responsibility, for which consent must be mandatory. Herd immunity relies on positive action from 95% of the population which means herd immunity relies on 95% of people consenting to vaccinations for their children. Herd immunity from measles is generally successful as a strategy in countries with an adequate supply of the vaccine, and with administrative structures in place to organise vaccination of children at the appropriate ages. Vaccines have been established as successful for decades, and their risks are known to be low. In recent years, countries France and the United Kingdom have experienced drops in the numbers of parents willing to consent to vaccination for their children, which has resulted in the UK falling below the immunisation levels required for herd immunity against measles, with a recent poll in France indicating that around 1/3 of adults believe vaccines are either ineffective or dangerous (though the rates of vaccination have not yet dropped by the same rate). Explanations for these failures in public health vary from place to place, and the WHO has explored the possible influence of religious views,¹⁵⁴ the rise of individual responsibility over social responsibility,¹⁵⁵ and of educational failures in the countries involved.¹⁵⁶ The 2019 Wellcome Gallup poll analysed worrying anti-vaccination views held increasingly widely in France and suggested that localised mistrust in the government might be one factor to help explain the results.¹⁵⁷

Third Dialogue: Emotions and Medical Advice

Localised reactions of mistrust following notable public mistakes by governments partially explain a more general trend that has seen the rise of anti-vaccination beliefs expressed in jurisdictions which have had vaccine strategies for decades, but do not entirely explain the change. Competent adults who refuse consent for established and well-evidenced procedures on behalf of their children are at liberty to do so in many global jurisdictions, particularly liberal democracies where social convention and interpersonal responsibility is the main reason most people comply with vaccine programmes.

In clinical settings, parents are offered vaccines for their children at appropriate age markers, with an explanation of the risks of not vaccinating for the health of the child and the herd protection vaccines afford the children of others. The ethics of vaccine mandates, *i.e.*, to enforce vaccination against patient autonomy, where the success of procedures like vaccination are well established, does vary with national politics, but remains controversial in most democracies. Proponents of legal enforcement of social responsibilities including vaccines suggest that it is justifiable for governments to impose fines and other severe

penalties on adults who refuse vaccination for themselves or their children¹⁵⁸ classing vaccination as a responsibility alike to adherence to tax regulation. Similar arguments when used in relation to organ donation are also extremely controversial, even though the donor body can no longer make biological use of the tissues as in life.¹⁵⁹

Legal intervention to intervene on bodies without consent can be justified in cases where it has been established that individuals do not have the right or capacity to decide for themselves. Where the effect of a choice is not only, or not primarily, likely to damage their own health but would potentially damage the health of others, enforcement can be within the mandate of a governing body. However, where the effect of the action of individuals would only damage their own health, there seems to be little justification for enforcement, for example to compel an adult who refuses to consent to life-saving treatment for themselves or others. Proponents of enforcement and paternalism in matters of public health have argued that certain unconscious biases lead to bad reasoning about vaccines, particularly in assessing risk, and so lead to patients irrationally refusing consent which they ought to give, as rational consequences of their values. Irrational choices are still not generally accepted as reason enough for authorities to overrule the consent of capable adults, and it is generally accepted that people can be capable of irrational and poor decisions while still being entirely competent to decide for themselves.

So, it becomes important to ask of transhumanist interventions whether harm would be done only to individuals who choose for themselves, or whether any harm associated with enhancement could be called a public health issue.

5. When emotions guide public health decisions are their reasons less valuable than other kinds of reasons? For example, disgust about face and eye transplants might lead to fewer donations of corneas.
6. Are governments more justified in intervening to offset emotional biases which could reduce public health responsibility, if such reasons are less valuable?

To answer these questions satisfactorily, it is important first to know whether allowing bias that might prevent public health successes is ethically problematic for public health authorities. If it is negligent or irresponsible to allow oneself to be guided by bias or irrational

feeling, it seems more acceptable that certain decisions should be addressed, if libertarian intervention were possible. Unconscious bias can be overcome by deliberate effort and reflective self-inspection of motives and desires in many cases, particularly when the need to reflect is made apparent by an external force. Correspondingly, in cases where nudges seem most effective, they act by relying on known biases, or by identifying and so diminishing bias. Common biases that can be used by nudge theory include the tendency to favour the status quo, or to judge situations in terms of their neatness and balance. A change in procedure or legislation, which would make decisions fit the opt-out (not opt-in) choice architecture, would make use of a general rational bias that favours of the status quo.

Responses to transhumanism are not static but should also be understood in terms of their shifting suitability within cultural frameworks. It is central to modern medical ideals that patients should be competent to make autonomous judgements about their own treatment. Bodily autonomy in medicine, if it is to promote beneficial outcomes and justice, relies on individuals taking personal responsibility for living well and making appropriate use of healthcare services, where they have influence to do so. The basic availability of healthcare, as well as education, providing a threshold for healthy lifestyle for all citizens, reinforces the general expectation that autonomous individuals with means and a basic understanding of healthy living should tend to value their own longer health-spans. A public responsibility to promote longer health-spans is only conceivable as a real group project because of important technological and scientific shifts in normal science, which might allow core assumptions about healthy lifespans to change.

Emotion heightens awareness and changes the epistemological landscape as it appears to the individual. It seems improbable that the influence of fear, hope, or even anger could be kept away from medical decisions or consent-giving altogether and consent is improved when the experience of emotion allows for full apprehension of risk.

As emotions can improve rational decision-making, and often motivate behaviour more effectively than factual knowledge or understanding, attempts to change behaviour are more likely to succeed where they do not ignore emotions.

Regarding many personal decisions, deliberate interventions to change the emotional response of individuals to stimuli seem, at best, patronising, at worst, manipulative. The use of emotions in any intervention made by governing bodies tends to create unease, and the unease that individuals feel about emotional persuasion by governments may help draw a distinction between persuasion and propaganda.

Implicit in the aims of the transhumanist movement is the premise that human self-control and progressive change of human capacity is for the best. In the existing literature, and in practice, this premise raises questions about when the relevant biomedical or existential control is best for the individual, for society, for subsequent generations of human beings, and for long-lasting norms of judgement and behaviour. One significant consideration underlying discussion of the ethical good of the transhumanist project, then, is the positive role of emotion in self-control, beliefs, desires, behaviour, and meaningful life.

Which emotions are relevant to transhumanism as human enhancement medicine? Why?

Some painful emotions, like boredom or fear, can be indicators that a situation is not good for us. We act on judgements that a situation either not good for us, or not good as construed by our normative judgements. These judgements might also be fast, automatic, heuristic, or biased. They do not interfere with our rational judgements, rather we might wish to explain that such emotions contribute to the creation and maintenance of reason by focusing our awareness and motivating behaviour. In this final dialogue, I explore the legitimate role of emotion in reason, as another element of the internal processes contributing to autonomous choice. Alongside bias, emotional involvement, for example about medical judgments, might not seem to lead to 'cool' rational decision making and cannot very often be avoided.

“Our angers, envies, and fears have been accused of interfering with proper reasoning, to favour irrational behaviour, and to elicit immoral actions.” But many modern theorists suggest that “emotions are both necessary to the proper functioning of theoretical and practical rationality, and essential to moral action. In brief, emotions would allow us to think and act more appropriately, both from the point of view of prudence and ethics”.

Importantly for the thesis, it seems possible that certain emotional responses should be beyond the reach of ethical nudges if these nudges should not interfere with reasoned preferences. Interference with reasoned preferences risks interference with autonomy and could constitute propaganda if used by governments.

Emotions and self-improvement

In the earlier discussion of medical autonomy, it was clear that all manner of external influences can cause us to change our minds, and that not all such influences were a threat to autonomy. Primarily, when agents encounter new relevant information or when circumstances seem to change, external influences prompt appropriate and useful rational reflection and reconsideration. For example, when a clinician presents a patient with the likely outcomes of a desired medical procedure, the new information about risk should cause reflection and reconsideration and might result in the loss of desire for the procedure. This kind of external influence seems much more likely to enhance autonomy than to undermine it, making the patient better able to decide, which is the foundation for the medical requirement of informed patient consent.

One problem, given the role of external influences in shaping behaviour and preferences, was bias. Biases like framing and presentation effects were not similarly beneficial to individual autonomy. For example, the different behaviour of a patient presented with the 'risk of death', to one presented with the same risk as 'chance of survival'¹⁶⁰ suggests not all similarly informative prompts to reconsider a decision are similarly beneficial. Indeed, for any apparently autonomous behaviour, belief, or desire, no matter how controlled or personal it might feel, behavioural science since the 1970s seems to have identified possible hidden influences that *actually* drive behaviour. Heuristic short-cuts and unconscious cognitive biases are not preferences, beliefs, or desires however, though they seem to change the epistemological landscape of the agent in a similar way. Unconscious bias and heuristic short-cuts do change and drive behaviour, but do not present new information to the agent such that they can reflect. Implicit bias and heuristic short cuts also seem to be useful, though agents are often unaware of their influence, and not in a position to reflect on their own biases at the time of action. Behavioural psychology, since the 1970s has explained the presence of heuristic 'quick thinking' and more recent work (for instance by Kahneman) makes more of the usefulness of System 1, as it works alongside the slower and controlled, reflective System 2.¹⁶¹

When psychologists identify the effect of bias, or unconscious quick thinking on behaviour, they appear to highlight a lack of accurate introspection available to agents, even when they reflect. For example, implicit bias can guide behaviour without appearing in the account given by the individual of their own reasons and preferences.¹⁶² In such cases, participants are shown to deny the involvement of bias, even when presented with evidence.¹⁶³ When

Cass Sunstein explores the role of heuristic and implicit bias in the uptake of progressive actions and behaviours, he made the distinction between those individuals who masked their true beliefs, behaving as if they behaved how they imagined society would prefer, and those individuals who masked nothing, and behaved according to social expectations. The distinction is important in terms of what can be expected when deliberate interventions aim to change the direction of social influence, as those masking their views will change behaviour but need not have changed their minds. If the benefit of a libertarian paternalist intervention is that it changes minds, and therefore behaviour, rather than forcing behaviour on those who disagree, then such belief masking, as a reaction to strong social norms, is a complication.

In terms of a medical decision, imagine the following scenarios, where a nudge might influence behaviour:

- i. An individual has strong beliefs, but masks them, because they feel their beliefs are not shared by society and would subject them to condemnation. An intervention makes their belief more socially acceptable by changing the social norm. The individual changes their behaviour. Their beliefs are unchanged but are unmasked.
- ii. An individual has strong beliefs that go against the grain of the social norm and their behaviour, which is the result of their beliefs, exposes them to condemnation. An intervention makes their belief more socially acceptable by changing the social norm. The individual does not change their behaviour or their beliefs, but now is less likely to be condemned for acting or expressing their beliefs.
- iii. An individual has strong beliefs that are compatible with the social norm. They act on these beliefs and receive social approval. An intervention makes their beliefs less socially acceptable by changing the social norm. The individual comes to change their mind (change their beliefs) and so changes their behaviour, or else their behaviour changes and they are led to change their beliefs, on reflection about their behavioural change. They continue to receive approval for their behaviour from others in the same position.
- iv. An individual has strong beliefs that are compatible with the social norm. They act on these beliefs and receive approval. An intervention makes their belief less socially

acceptable. The individual changes their behaviour but does not change their beliefs, which are now masked. They continue to receive approval.

- v. An individual has strong beliefs that are compatible with the social norm. They act on these beliefs and receive approval. An intervention makes their beliefs less socially acceptable. The individual does not change their beliefs or their behaviour, which exposes them to condemnation.

From a public health standpoint, it might not matter whether individuals mask their beliefs or not, as long as they engage in the preferred behaviour. For example, whether or not individuals believe vaccines work, whether or not they would publicly endorse vaccination, if they vaccinate their children then herd immunity will remain. However, as recent research from the WHO demonstrates, erosion of trust in even established medical practices like vaccination can happen if enough individuals do not believe vaccines work and start to feel and behave differently. The Wellcome Global Monitor discovered that around one third of adults did not believe that vaccines were safe or effective, when polled in 2019. Though this finding has not yet been matched by a corresponding dip in vaccinations in France, it certainly was registered as a cause for concern for global health, signalling a more fundamental lack of trust in the health service, and in government public health advice.

The transformative force of social condemnation and peer approval is not underestimated by policymakers who rely on the global investigations carried out by, for example, the Wellcome monitor. It is recognised that the dip in trust of vaccines in France might risk becoming a dip in desirable behaviour, if familiarity with anti-vaccination arguments lessened the social condemnation attached to public association with the belief. In such a case, the aim of policy will aim to reinforce the existing social norm, which appears still to be supported by two thirds of the adult population. It is still possible, though, that unpopular anti-vaccine beliefs will simply be masked.

It is widely accepted that emotional states like fear or anger can influence desires and beliefs, as well as individual susceptibility to apparently more complex and socialised feelings of shame or moral disgust. Increasingly prevalent in the literature on emotion is the suggestion that emotional involvement in behaviour supports rationality, instead of undermining it.

The challenge for policymakers, aiming for the ideal third scenario, is to avoid putting too many people in a position where they must mask their strong beliefs to avoid condemnation,

whilst still making an impact on the social norms such that undesirable behaviour does not go unchecked.

It seems clear that emotions like fear, excitement, anger, guilt, boredom, and others also motivate and change behaviour relevant to public health and might do so in ways that are not identical with 'clear headed' rational deliberation. However, unlike implicit biases or heuristic 'quick thinking', emotions tend to motivate and influence by increasing awareness and focus on certain facts and ideas over others, rather than by bypassing awareness altogether.¹⁶⁴

Often, emotional involvement in behaviour might pass below conscious awareness, for example as the influence of a certain mood be something about which an individual is generally but not specifically aware. For example, we would expect an individual to be aware of their anger, or able to identify it in later reflection, but the same might not be possible, to anticipate the influence on behaviour of a more habitual or pervasive aggressive mood.

Distinct from moods, though, emotional involvement in action more usually is the cognitive evaluation of stimuli, resulting in different epistemological weighing of beliefs. Emotions are consciously available to the agent and are more usually compatible with reported beliefs. *i.e.*, I fear proximity to tigers because they might bite me. The emotion is reasons-responsive, although it might sometimes 'get the better of me'. On visiting the zoo, fear of the tiger might rationally be tempered by a belief that the enclosure is secure. This does not mean a sudden roar will have no effect on my composure, and such a temporary loss of composure does not indicate my rational ability is diminished.

A cognitive account of the emotions suggests that the revised epistemological landscape during experience of the emotion can be rational or irrational, depending on the content and appropriateness of beliefs and desires involved. Examples like the roar of the caged tiger also seem to demonstrate that temporary lapses in strength of belief because of an emotion do not automatically mean an irrational individual.

Problems for the cognitive account include the challenge of describing the fit of beliefs and desires into the affective realm. In brief, some accounts suggest that emotions simply *are* beliefs or desires, but this seems much too reductive, missing other parts of the experience of an emotion, such as the physiological experience of bodily changes. Alternatively, 'umbrella' accounts suggest emotions are such that they can be treated as a belief or desire, when explaining how the emotions figures as part of production and control of behaviour.

So, an emotion could be called 'belief-like' or 'desire-like'. Finally, the nomological view suggests that emotions 'carry beliefs and desires with them'. For example, anger is a bad feeling, an evaluative belief that there has been an offense, the belief that someone is to blame, and the desire to punish them. This would explain the link between emotion and behaviour by reference to truth conditionals about the emotional state, for example 'If I am angry at x, I believe x to be responsible for an offense'.

In the thesis, I will assume a cognitivist account of emotion is correct and suggest that emotion is important for individuals to arrive at rational decisions, rather than being an inconvenience that detracts from reason. I refer to an argument from Sabine Roeser to identify why it is more beneficial and justifiable to engage emotionally than to use informatively persuasive interventions but aim to avoid contrasting emotional persuasion with rational persuasion.

Emotions do change the decisions we make, and their effects are well documented in cases like judging risk and judging value. However, differences in rational decisions caused by different emotions do not mean emotions undermine rational abilities, as rationalists or subjectivists about emotion¹⁶⁵ suggest. A cognitive view of emotions allows for the idea that emotions are perceptions of reality, rather than being arbitrary or socially constructed projections.¹⁶⁶ Emotions as perceptions, although like any perceptions they are fallible, allow us to assess 'what is really there'. Emotions allow us to judge what seems to be 'really there' and to perceive moral reality.¹⁶⁷

As far as fallibility is concerned, moral mistakes can be made where situational emotional responses are inconsistent with what an individual expects or believes in general. As an example, when individuals feel fear or unease about the use of a particular technology or innovation, like genetic modification of foods, they will prefer not to eat GMO foods, despite recognising that GMO development is likely to be generally beneficial, and unlikely to be harmful. The contradiction, of accepting that GM foods are likely to be beneficial in reducing food shortages and to reduce harm by controlling crop disease, while refusing to eat GM foods because of fear or unease, creates what Roeser calls a NIMBY problem, as a kind of free-rider problem¹⁶⁸: "it's ok for everyone else but not for me". Inconsistencies like these should be reduced by deeper reflection on the emotion (revealing the inconsistency) but are also likely to diminish with feelings like sympathy or desire for cooperation.¹⁶⁹

Roeser also identifies situations where the role of emotion is not to cause moral mistakes (like NIMBY cases), but rather to lead to poor evaluation of empirical evidence, or 'blindness

to descriptive facts.¹⁷⁰ In situations like these, presenting the agent with more information should be enough to offset the error. Public health policy aiming to influence emotions instead of to engage with people persuasively using facts might appear unwarranted in many situations, precisely because presenting information and allowing agents to re-evaluate for themselves appears to offer a lighter touch and might be more effective. Roeser suggests that the technocratic approach might lead to 'complexity neglect' if it aimed to address with emotions the feelings of fear or unease toward technology that were not like the NIMBY cases.¹⁷¹ However, it is not always possible to present the layperson with descriptive facts that will sway their evaluations. In the next section, I will discuss the idea that, even where emotion does not cause moral error, like the NIMBY or free-rider cases, emotional intervention is sometimes acceptable and beneficial.

On influencing emotional preferences

Certain matters elicit emotional responses which are not directly the result of evaluating risk or benefit of outcomes, but not all these cases are contradictory or irrational, like the NIMBY or free-rider problems identified by Roeser. In cases where there is no clear contradiction in the emotional influence on choice, but the emotion is not 'useful' in directing the chooser towards the best outcome, it is not always possible to influence choice only by presenting new or better descriptive facts. In these cases, it appears to be beneficial to influence choice using emotion but as identified by Roeser, such intervention appears technocratic, and further than Roeser's claim, might risk becoming technocratic propaganda.

Although the thesis will not adopt his non-cognitivist approach to the bodily experience of emotions, William James also develops a useful framework for being able to distinguish between the genuine choice of non-empirical beliefs available to agents, and those which are beyond persuasion. In his *Will to Believe*,¹⁷² James' argues that there are situations where an individual *can* decide what to believe, particularly in matters where empirical evidence is no help, for example in the matter of religious belief, or belief in God. The ability to choose belief in this way relies on the idea of critical self-reflection and deliberate self-transformation, capacities which will appeal to the transhumanist movement, in its approach to public health responsibility. James' framework does suggest that preferences, emotions, and the non-empirical have bearing on how beliefs and desires can be evaluated and formed. James intends to identify cases where belief can be governed by the will, and to explain why in some cases agents are in control such that they can choose to believe at will, but that this

control is not mere 'wishful thinking' or self-deception. Choices between hypotheses which motivate action and drive behaviour are not always alike and, in involving James' framework, I identify some beliefs and desires which are the result of emotions and intuitions because they are culturally alive, exploring the appropriateness of external interventions that target them.

Important choices between hypotheses are those where the relevant ideas are alive to the chooser. Similarly, it is more significant to choose if there is no way to avoid making a choice, and self-transformation is more likely if the effect of a choice is momentous for the chooser. For James, if a choice between hypotheses is the subject of a will to believe, it should be *alive, unavoidable, and momentous*. Those choices that imminently confront the agent will therefore occur in a context they understand, about a decision the agent cannot avoid, and where opportunities to make the choice are not ten-a-penny everyday occurrences. Such choices can be transformative in a way that distant or unfamiliar, avoidable, and trivial choices tend not to be. It is these self-transformative choices that should be the subject of more careful observation, and perhaps are also more attractive targets for intervention by governments, because of their capacity to cause or avoid harms, and to change the population for the better.

Preferring a certain flavour of ice cream over another, holding this preference in mind, and using it as a reason to guide behaviour, relies on cultural norms and contingencies to make the positive feeling in favour of one option relevant to behaviour. Cultures share particular ideas about the reasons to pick snack foods, i.e., choices not based on nutrition, but on pleasure. The range of normal flavours varies, such that even in the context of soups, my favourite flavour is broccoli, I should know it does not apply to ice cream choices. In some places there might be no ice cream at all, or no chocolate, and this changes what it is likely that I will prefer, or whether I can be asked if I prefer at all.

In the ice cream case, given a preference for chocolate over vanilla, any choice between them is also entirely avoidable, given the possibility not to have ice cream at all, or perhaps to choose strawberry, or to have both chocolate and vanilla. It is highly unlikely that this choice will be the last of its kind for me, nor is it likely that choosing one flavour over the other on this occasion will directly change the course of my life. For cases like this, therefore, preferences and their effect on behaviour seem to be impractical and undesirable as targets of intervention by governments. Any paternalist intervention intended to change ice cream

flavour preferences and therefore to change behaviour, even if it were carried out with beneficent motives, appears entirely excessive and unjustifiable.

If a hypothesis is alive, it is intelligible to the agent because it “appeals as a real possibility to him to whom it is proposed”.¹⁷³ This means that, given a live choice between hypotheses, neither hypothesis should seem alien or entirely outlandish. One example given by James is that a person might choose whether or not to have faith in the religious tradition of one’s own society, that either choice is plausible in the way that choosing to adopt the religious tradition of another society is not, because unfamiliar traditions would not make an “electric connection with [one’s] nature”.¹⁷⁴ Desires, similarly, are only alive in a society with knowledge of the desired item, and given some kind of shared understanding about what it is to desire something. One current example of a choice between live hypotheses might be whether the everyday reduction of car-use has a significant impact on reducing the human contribution to climate change. Because it seems clear that reducing emissions is important, the hypothesis is alive in our society in a way it might not have been for previous generations. However, because the evidence also suggests that industrial waste and corporate emissions are a much larger contributor to climate change, it is also a live hypothesis that individual car journeys are not importantly contributing to the problem. At the current time, both hypotheses seem live, but it is foreseeable that this will change with evidence. In the absence of evidence, a prudential approach to making this decision for oneself seems to be an emotional one, and is likely to be driven by various preferences and practicalities, a host of other commitments and beliefs about appropriate behaviour, and perhaps even self-identification as, for example, ‘green’ or ‘a cyclist’.

A forced preference cannot be sidestepped in the way a choice between ice cream flavours can be avoided. Forced choices might include whether to drive to work or not, or whether to try to buy healthy groceries. An agent in the supermarket must either buy items that are nutritious, or not, and there is no way to avoid this choice. This is true even if the agent is not aware of the choice their behaviour commits them to, because even an agent who buys nothing must act one way or the other, either having bought healthy food, or not.

Importantly, if on one occasion someone happens to buy healthy food, this does not mean they always will. Similarly, if on one occasion they do not buy healthy food this does not immediately dissolve health or necessarily undermine their apparent understanding of healthy eating. Beliefs and desires that make behaviour more likely to be habitual are more important than instances of behaviour, therefore. It is the choice of the hypothesis, about

healthy eating, rather than the simple choice of actions that can be momentous and self-transformative.¹⁷⁵ Over time chance actions *might* lead to a momentous shift, transforming the individual from someone healthy to someone unhealthy, but this is not enough to rely on for public health policy. Momentous hypotheses are those which, once chosen, are not easily reversible, or are transformative for the individual, and these seem the most fruitful areas for policy to intervene.

If every choice between hypotheses is bound to be (i) alive or dead (ii) forced or avoidable (iii) trivial or non-trivial, choices can be organised by these differences which would separate the forced trivial from the forced non-trivial, for example, and explain why different treatment of choices might be justified. Based on James' framework, it is possible to divide choices between eight permutations which allow discussion in terms of general salience and urgency to populations, within which emotions as reasons, given a more modern understanding, act to change circumstantial salience and urgency at key moments.

As James' argument is not about the use of empirical evidence to make decisions (James' original discussion is about belief in God, about which, he takes it there can be no empirical evidence to settle the matter in either direction), the role of factual information which might otherwise be involved in some of the examples below will briefly become inconsequential. Whether hypotheses are alive is in part determined by their fitness according to current paradigms familiar to the agent. The agent may be aware of healthy eating but find that learning about how to keep healthy does not have an impact on their desire for fatty foods. As discussed in the previous section, it seems likely both that the agent's emotional responses are evaluated differently because of evidence presented to them by society, and that evidence is weighed differently because of emotion. In the cases below, it will be assumed that emotion is the overwhelming motivation for behaviour, though the hypotheses are both alive, in the sense of being sensible to the agent in terms of the evidence involved. Conversely, dead hypotheses are already, by definition, those which the individual is systematically discouraged from taking seriously just because of existing social norms, expectations, and the socially normal worldview supported by available science. In the four permutations of James' conditions where both hypotheses are alive, therefore, there seem to be good (socially supported) reasons to think there is a real choice between two possible hypotheses, neither of which would be entirely alien or inconceivable to the chooser, and I will not discuss interventions that would aim to change preferences about dead hypotheses.

Alive, forced, trivial choices: The beliefs and desires that inform the choice of various grocery items effect public health outcomes. This is a real cultural concern, which might seem to justify intervention. Beliefs about the extent to which choices change health, or that an individual feels they are in a position to control their health, also involve personal preferences that might be socially transformative over time for individuals. Because healthy eating is a culturally salient choice, and because agents must choose how they behave as regards their own healthy eating, if they choose inaction, outcomes of trivial choices can eventually be non-trivial. However, individual instances of buying groceries are not the same as the preferences involved in making individual choices within healthy food groups, intervention to try to change one is not the same as intervention to change the other.

Alive, avoidable, trivial choices: As above, a choice might be alive to the agent, and relatively trivial (because ice cream preference is hardly likely to redirect the course of a life, or to be a once in a lifetime event) and might also be perfectly avoidable. The choice between chocolate and vanilla ice cream might never come up for some agents, and even if it does, there is always the choice to have no ice cream at all, or to choose a different flavour. The impact of this kind of decision is so minimal that any kind of intervention to change behaviour also seems extremely unnecessary, but it is not clear that there would be no similar cases where impact of the preference caused harm that did justify intervention. If, by the previous definition of a trivial choice, 'alive, avoidable, trivial' choices would include antisocial actions like littering, then something seems amiss. A trivial decision, as well as being one which does not alter the course of the life of the one who chooses it should also not cause harm to others. To clarify this position, I will describe 'non-trivial' instead of 'momentous' choices in the rest of the examples, because the choice to avoid harming another person may not in itself be momentous, though it is non-trivial whether or not antisocial behaviour or harm is chosen. The clarification, that trivial choices must not cause harm, would make irresponsible or antisocial actions like littering 'alive, avoidable, and non-trivial' (about which choices, more below).

Alive, forced, non-trivial choices: For James, this combination was a choice most likely to be self-transformative, and the choice where it is most likely that beliefs (and desires) could be modifiable at will. For the purpose of a discussion about when and whether governments are justified in making interventions to direct preferences, to guide behaviour, minimal intervention is preferred. To minimise intervention, it is best for governments only to intervene where small interventions would be most likely to have large positive impact, so

the best candidates are occasions where individuals must choose, and the effects of their choice will be lasting.

In the case of organ donation, the nudge implemented was to change registration from opt-in to opt-out. A change in the environment intends to change the way society feels

Alive, avoidable, non-trivial choices: These choices are much more serious for the agent than preferring certain ice creams, or preferring not to litter, but not every non-trivial choice between two hypotheses includes the implicit choice resulting from inactivity. Directing forced choices, those where only two choices are available, is likely to be less complicated to implement with minimal force than attempts to direct multi-faceted avoidable choices. A concern for behavioural insights in particular, and policies of intervention in general is the risk of unforeseen consequences, so simplicity is important. In the case of organ donation, an intervention intended to increase the number of registered donors by changing the default is unlikely to have direct behavioural impact wider than the decision to be a donor. An intervention to nudge people toward a more generalised belief that medical professionals are knowledgeable could have a number of consequences, which are harder to predict.

The relevance to transhumanism

The argument of the thesis does not rely on asking whether the affective responses people experience in response to enhancement are enough to support that we, as a species, should pursue transhumanist goals, nor whether they are enough to judge that transhumanism is an ethical endeavour. Rather, I assume continued work on a transhumanist project of some kind is inevitable and suggest that careful management of important affective responses alongside more analytic and deliberative responses will determine how transhuman enhancement can become normal and produce flourishing life. Because the demands of maintaining autonomy are such that intervention into the beliefs, desires, actions of individuals should be minimal, and because the transhumanist paradigm is not yet settled enough to make risks and benefits certain, efforts to re-shape emotive debate should be cautious, but are necessary.

The sort of hypotheses transhumanism presents to the average individual now concern attitudes to new technology, to responsibility for personal and public health, and ways to interpret the value of life and of ageing. The extent to which individual hypotheses are alive to individuals varies by social setting, time, and place. For some, particularly those working

in the biotechnology industry, transhuman enhancements are currently much more 'alive' than for those outside the industry, or for those living in societies where biotechnology is not yet well established. Initially at least, transhumanism as optional enhancement is likely to involve the use of genetic information to guide treatment options. It is likely that there could be the expectation of choosing whether to engage with enhancement, in a way comparable with the paternalism of patients being free either to take medical advice or not, with all the same implications for personal responsibility when bad decisions are made against best advice.

Transhumanism seems to require changed attitudes to the application of enhancement technology to the perceived normal degeneration of human bodies. While fear of interventions might be natural or predictable given our neurology, it is also important to note that normal emotional responses can undermine as well as contributing to autonomy. True consent and autonomy sometimes rely on automatic fear responses operating automatically but are sometimes undermined by strong emotion. For transhumanism to justify interventions that would seek to go beyond familiar medical ground, change should not be forced on the unwilling, for ethical reasons that follow from the preliminary discussion of autonomy and consent earlier in this chapter, but force might not be effective in such cases either.

For ethical transhumanism, the success of which will rely heavily on certain emotional shifts occurring in the general population during public debate and engagement with emerging technologies, there seem to be emotions that can make bodily enhancement and personal change appealing, or that make alienation or existential undermining from bodily change more likely.

So, there seem to be two significant questions about the emotional impact of enhancement as public health, for transhumanism to answer:

7. Can anyone control these emotions? Do individuals decide how they feel about enhancement?

8. Should anyone attempt to control these emotions?

If emotional responses support individuals giving informed medical consent, should their emotions be targeted by government influence intended to improve medical outcomes? Even if the result of interference is a population systematically more likely to take up beneficial enhancement opportunities, why might this not be acceptable? Are there appropriate situations where intervention would change personal preferences in favour of transhumanism?

SECTION TWO

4

Is Human Enhancement a Public Health Problem?

Are established public health strategies available to transhumanists to apply to a new transhumanist global health paradigm?

Introduction

Can advocates of widespread medical human enhancement use existing arguments in support of public health interventions? Yes, they can, because human enhancement medicine aims to strengthen generally accepted public goods; because not all public health measures are limited to those pursuing publicly agreed goods, particularly in cases where intervention does not restrict freedom; and because the outcomes of medical human enhancement do not seem likely to produce outcomes that fall beyond general understanding or appreciation of bodily wellbeing.

Critics of medical human enhancement, for example those opposed to the use of gene editing to produce and expand certain desirable traits, suggest that there is no justification for imposing or encouraging enhancement, particularly if people do not value enhancement. Because transhuman principles are not generally accepted and propose radical change, any attempt to rely on democratic principles of public responsibility fail. Critics of genetic editing for enhancement suggest that autonomy is lost when people are enhanced pre-natally by their parents, and that deliberately increasing the homogeneity of human traits in this way would prove morally damaging for the species, in the absence of a medical need. Critics of transhumanism suggest that people are right to be afraid of the implications of medical human enhancement, and of enhancement applications of medical treatments not only because of the physical changes they produce but because of the leap needed against democratically agreed existing practices. A further problem is that attempting to reduce fear by persuasion in the absence of generally established public feeling in favour of transhumanism seems inherently coercive, and amount to unjustifiable propaganda. Unjustifiable propaganda, loss of autonomy, and forced homogeneity would be ethically damaging to societies that used such tactics.

In this chapter, I will outline three ethical strategies that are commonly used to evaluate the use of authority in medicine, and to assess the justifications for authoritative interventions in

public health, which might also seem to aim at homogeneity, using control, or persuasion, for example by aiming for consistent results across diverse populations with social measures. I explain how each strategy of identification and justification of public health measures is applicable to the criticisms set against the prospect of ethical medical human enhancement. About each existing strategy, as they also set about to evaluate public health risks, I consider whether critics of preventative public health would find even more reason to resist human enhancement, and whether these criticisms can be overcome in genetic enhancement cases relating to negligible senescence and medicine geared towards life-extension. In response to each criticism of transhumanism and medical human enhancement, I argue that concerns about the use of persuasive measures in public health need not be heightened when the medical technology involved is emerging, rather than established. This is because the methodology of applied medical knowledge admits of ethical judgements that focus on narratives of social wellbeing, and the value of certain outcomes can be 'live' before individual clinical applications are supported by practice.

Background:

This thesis argues in support of nudges to bring about transhumanism as public health. This chapter will aim to assemble and defend a framework based on existing ethical strategies which marry general medicine and public health methods, to find what is acceptable in transhumanist public health nudging and what is not, particularly given the uncertainty of future enhancement applications of medical technologies.¹⁷⁶

I address the concern that public health oversteps its justifiable bounds if it restricts freedoms other than in times of acute emergency. Unlike outbreaks of disease, enhancement medicine often tackles expected causes of death in old age and the causes of ageing, and it does not seem realistic to describe the steady rate of deaths in old age as an emergency in the same way that the outbreak of infectious disease or the lack of clean drinking water might be described as an emergency justifying interventions and restrictions. I refer to public health strategies for planning and ethical justification of interventions, though there are differences between illnesses like heart disease or cancer and the processes of gradual slow ageing, as the cause of human suffering and lost life. It seems increasingly likely that both can be addressed by appropriate medical care given timely intervention, and that the suffering and increased risk of illness in ageing is an appropriate target for global public health work.

Public health strategies make appeals to the authority of the values and preferences generally held by people in a society, both by experts and the public (laypeople). In public health, it seems right to pursue public goods using resources that have been set aside for the maintenance of national wellbeing, and public health theorists have sometimes preferred to limit the definition of public health only to those interventions relevant to *generally agreed public goods*.¹⁷⁷ A typical example of a generally agreed public good is access to clean drinking water, where there is little or no disagreement about the value of the resource or the danger of its absence. A more culturally variable but similarly globally relevant example of a generally agreed public good is the prevention of the spread of harmful infectious diseases. There may be different infectious and dangerous diseases at different times in different parts of the world, and so there are likely to be different cultural norms guiding policy, but group action, prevention and vaccination are generally agreed good uses of resources when harmful infectious diseases seem likely to spread. Accounts of public health as generally agreed public goods would not include all risks and challenges that can shorten life, however, and they would not tend to include efforts to reduce rates of obesity or to curb gun violence. As a result, a public goods account of public health might seem even less likely to include heritable enhancements among public health measures.¹⁷⁸ Public health measures that improve the health of the general population, such as infant vaccination, are not identical to and do not usually share an ethical framework for reasonable interventions with more straightforwardly social measures, for example attempts to tackle gun violence. Further, even between apparently comparable social measures like gun violence and dangerous driving, or between apparently comparable public health measures like infant vaccination and organ donation, it is not always the case that the interventions accepted for one problem will be accepted for the other, nor that interventions will generate the same public response. The familiarity of the goals of interventions, and compatibility of the interventions with more general social and behavioural norms can be a deciding factor in the success of interventions, which might seem to predict the failure of transhumanism. Unfamiliar goals, for example the transhumanist goal of negligible senescence, particularly those that appear to be incompatible with social norms might seem to make the transhumanist project unlike widely accepted public health or even social policy measures. If public goods and general agreement on public goods can be construed as many public health theorists prefer, then transhumanism can be said to have more in common with vaccination programmes than it does with gun control.¹⁷⁹ The ethical framework supporting public health interventions increasingly already includes efforts to increase health-spans by encouraging bodily

interventions, though their targets are not strictly public goods in the same vein as is access to clean drinking water.

Transhumanism, insofar as it embodies longstanding goals and principles of human ethical activity in medicine, seeks to ascertain whether and why death in old age should be treated any differently than deaths characterised as preventable.¹⁸⁰

In this chapter I will discuss the way models of death risk have changed, and how this is likely to affect the proposed approach to transhumanist medicine. Health Adjusted Life Years (HALY), then Quality Adjusted Life Years (QALY) and Disability Adjusted Life Years (DALY) were developed to bring evaluative assumptions and processes more into the discussion about preventable and reasonable death, with the more general aim of modelling deaths in more and more systematic ways to allow just policymaking and just public health outcomes. At first, models aimed to increase detail¹⁸¹ and then to reduce the influence of hidden assumptions on reported morbidity.¹⁸² The final goal of modelling adjusted health span and life span of human populations is to understand with greater nuance why people die prematurely, without burdening the figures with prejudices that might prevent progress. The philosophical problem in the QALY literature is whether it is better to try to remove assumptions and evaluations from measures of morbidity and mortality that ultimately decide what we consider to be urgent or premature, or whether to acknowledge assumptions and evaluations as a transparent and systematic part of the measurement and use them as a guide to distributing public health resources. In the latter case, transhumanist negligible senescence would need to be better defensible as a generally agreed public good, and even in the former case, evaluations seem unlikely to vanish from the distribution of public health resources altogether.

Marrying public health efforts to improve the core capacities of human lives with coordinated campaigns to address the behaviour of individuals, behavioural insights is an increasingly used vehicle for governance in democratic societies. Nudges can be used by authorities to help public health and other kinds of public campaigns engage with people's emotions, and with their unreflective evaluative processes to achieve better outcomes without the need for restriction or coercion. Nudges in public health, used carefully, can lead to changes in decision-making that can lead to changes in habit, reflection, and in preference which are more likely to last. As a method of guiding people toward more beneficial rational choices, nudges should aim to promote goals and behaviour most people already value but have trouble committing to or thinking about in appropriate ways, for example whether to

start saving for retirement or committing to regular exercise with long-term goals in mind. In response to arguments that valid consent, privacy, and distributive justice are challenged by nudges in some settings (for example in workplaces)¹⁸³, I agree to some extent. For example, nudges can be used to destructive ends and can be turned to uses similar to propaganda, or for example that people are already mostly rational and able to commit to the things they value without intervention.¹⁸⁴ However, though nudges are imperfect, public health engagement to prepare for human enhancement remains urgent and important, and choice architecture always exists whether it is deliberate or accidental.¹⁸⁵ Choice architecture can be improved, so that it does less harm than accidental structures of pressure, and this is important in preparation for the spread of uncertain, novel, or existentially provocative innovations, like enhancement medicine.

In the introduction to the thesis, considering the literature underlying current ethical concerns about enforced enhancement, I accepted the assumption from general medicine that neither entirely libertarian nor paternalist treatment of patients is universally appropriate, that patients need the support of experts about some decisions, and that bioethics precludes unjust use of force or coercion by medical professionals. It is important to explain the relation of the ethical framework that maintains this balance, and its limitations, to counterpoints in public health and social work measures, where good practice affects larger numbers of people in a broader way. I refer to three strategies to establish a relationship between principles in clinical (bio)ethics, broader public policy, and global governance of healthcare. This relationship, I apply to the challenge of ethically implementing transhumanism, as transhumanism seems most likely to succeed as individual clinical decisions contributing to species-wide self-evolution as public health.

The first strategy explains how the definition of public health can include broad and uncertain projects like transhumanism via human medical enhancement, even if there is not yet consensus that medical enhancement is a necessary or universally desirable public good.

The second strategy explains how measures of reasonable and preventable death, for example what counts as death from old age, have shifted considerably in the last twenty years. In this section, I describe a general trend that has been observed in the literature on QALY, DALY, and HALY, wherein measures of preventable death in old age have not simply shifted to greater and greater life expectancies as technology has improved, rather have changed the paradigm that evaluates how death prevention is treated. I explain why this shift is significant for transhumanists who wish to consider the approach needed to bring

medical human enhancement into the range of normal medical options, rather than simply wishing to continue the trend of normalizing old age for some individuals.

The third strategy explains how nudges are particularly good public health interventions for cases where there is general but not well-established consensus about what is good. Nudges are useful in the face of uncertainty and can allow trends of behaviour and preference to flourish: they can help ideas become beneficial habits. As widespread measures that take account of emotion and individual difference, nudges do not attempt to create discrete uniformity of action, belief, or lifestyle. The attempt to change general trends mirrors the approach of most public health, but not the regulations governing individual patient-clinician interactions, where uniformity and procedure are important safeguarding measures against unjust treatment of patients or unjust influence of individual clinicians. I briefly describe how nudges are useful in a public health setting and explain why this approach will also be useful to transhumanists who wish to reduce the effect of fear on public health decisions relating to human enhancement medicine, in the near future.

Public health as more than the provision of generally agreed public goods.

The account of public health as provision of universally agreed public goods, like clean drinking water, is too limited, given the increasingly social and technological means by which medical care is provided in much of the world. The principle of general agreement about what public health should aim to do, and which are the most urgent problems in need of intervention is sound, but it is insufficient to describe public health. In this section I agree that generally agreed public goods are one useful way to talk about the limits and justifications of public health intervention by governments but argue transhumanist projects should not be excluded on this basis, even though it can be useful to limit public health in this way.

Public health can and should involve more than the provision of clean water and preventative measures against infection, without using coercion or paternalism against individuals. A good example can be found in the recent case of organ donation policy in the United Kingdom, where social and legislative measures have made marked improvements in public health for transplant recipients through the NHS, in the last decade. The justifications required for making public health interventions of any kind should not lose their strength when public health responsibility is broad enough to address some social injustices. When social injustices

lead to widely varying health-spans within a population of citizens, they are also public health injustices. Projects aiming at generational and intergenerational self-evolution of species, like Transhumanism as human enhancement medicine, cannot be matters of individually motivated choice and must be public, as fair as possible, and global if they are to succeed.

Following a more recent trend in the literature that aims to widen the definition of health beyond the provision of generally agreed public goods, I resist some consequences of the more conservative premise that public health interventions ought to be those, and only those, concerned with the provision of generally agreed public goods.¹⁸⁶ In part, this is a more popular position in reference to therapeutic medicine, given the contemporary modern context of discussion about public health measures, which increasingly has included understanding of and proactivity about social injustice as a source of poor health.

Definitions that would limit public health to generally agreed upon public goods are cautious out of a reasonable concern that too broad a definition might weaken public health as a category of interventions, undermining the unique and privileged justifiable avenues of action available to public health policies. Public health policy sometimes relies on restriction of freedom and in democratic societies it is generally agreed that restriction of freedom should be minimized, except where freedom to act would cause unjust harm to others. In the context of government interventions to restrict freedom in medical decisions, paternalism is not generally considered desirable, as it has the potential to undermine the ethical principle of free and informed consent, which capable patients must be able to exercise in clinical settings. A more conservative view of public health, the limits of public health as a justification for interventions made by authorities (*i.e.*, by governments), argues that the inclusion of too many policies under the umbrella of public health would lead to paternalism which could fall short of the ethical principles required of just healthcare. Opposition, for example from Dworkin¹⁸⁷, includes the problem that some progress is not available to consenting individuals, and must be achieved by coordinated group action when there is group consensus.

Anomaly positions Dworkin's counterargument as a contribution to the public goods account of public health, as a reason to allow some apparently paternalist interventions to pursue public goods. Dworkin argues that there are some kinds of collective coordinated action which cannot be achieved by individuals choosing individually. This kind of group action might also require action taken voluntarily by cooperative individuals, but also relies on coordination by an authority so the efforts of individuals are not wasted. The most prominent

modern example of this kind is action on climate change. As in Dworkin's description of these special cases, individuals might benefit individually from ignoring the problem though the group would benefit most from general adherence to green initiatives. Occasionally in general discourse, there arises a disagreement about whether the contributions of individuals are significant enough, compared to the contributions of multi-national corporations, to be bound by such stringent demands to act against climate change. Individuals might benefit from ignoring green initiatives, for example: using cheap air travel; avoiding the job of separating their recycling from landfill waste at home; purchasing unsustainable, cheap and disposable goods like 'fast fashion' or plastic drinking straws, products containing palm oil, and so on. It is currently thought that, to help reduce climate change, all the above should be avoided by individuals, though the contribution of multi-national corporations would be noticeably greater.

Motivating individuals to take their own behaviour seriously, in global terms, is one problem for action against climate change, therefore, and the free-rider problem is another. If, as Dworkin suggests, the "immediate interest of each individual is furthered by his violating the rule when others adhere to it" and the case for intervention is that individuals nonetheless agree that the rule is beneficial, then there is sufficient reason for a law that restricts some freedom, and which need not be called paternalism.¹⁸⁸

Even proponents of much wider definitions of public health, those which would allow for the wider justification of interventions on the grounds of public health benefits,¹⁸⁹ acknowledge the importance of a defensibly clear and discrete realm of public health which should not be expanded indefinitely. The reason for this concern about a slippery slope of public health spreading is that public health might then include policy decisions about war, terrorism, and the penal system, which seem too far outside the intended purpose and privilege of healthcare interventions. Whether public health should include vulnerability in terms of predisposition to disease or untimely death does not seem challenging, in the same way as the inclusion of war, terrorism, and the penal system. Even the goal of extreme longevity does not appear to fall beyond the scope of existing policy and seems at least to have relevance to the just distribution of public health resources in the same way. Medical science allows greater understanding of the reasons for disease, the reasons for disparities of vulnerability, to morbidity and mortality, within populations, and it is accepted that changing best scientific theory should also guide policy, once there is sufficient consensus, however consensus about extended lifespans and about the meaning of an untimely death is missing.

Consensus is appealing when communicating scientific findings to the larger academic community, to the press and to the public, particularly when the findings have implications for applications in medicine, climate change, or social care. If, as Miriam Solomon suggests, scientists who ‘want to be heard in a public context’ would do better to ‘advertise their agreements rather than their disagreements’¹⁹⁰ then consensus at the point where the findings of medical science are communicated to the public appears to be the ideal. Consensus is not always possible, however, and Solomon describes an example of a case where consensus cannot be forged, or disagreement ‘managed’ to present a stronger public front for the science behind medical practice. In the example discussed by Solomon, a longstanding disagreement about the use of routine mammograms in women aged 40-49 is described as follows:

‘As of July 2014 in the US, the American Cancer Society, the National Cancer Institute, the Society for Breast Imaging, and the American College of Radiologists all recommend annual screening starting at age 40, while the USPSTF and the American College of Physicians recommend against routine mammography for women aged 40–49.’¹⁹¹

A longstanding lack of consensus about the use of mammography for women aged 40-49 is the result of a disagreement in theory and conflicting, insufficient, and inconclusive results of studies. The results of some studies suggest that screening and biopsy of suspicious lesions in breast tissue contribute to the growth and spread of breast cancers, increasing the incidence of cancer in women who undergo regular screenings.¹⁹² Other studies have dismissed these results as a ‘statistical fluke’, supporting instead the prior hypothesis that earlier detection of cancers results in better patient outcomes.¹⁹³ Further risks involved in controversy over the disagreement is the psychological risk to patients from the stress of screenings and the many false positives that are recorded every year, as well as the bodily harm caused by the treatment of cancers which might never otherwise have been noticed. In this disagreement, and other rare occasions when consensus has not been reached at NIH Consensus Development Conferences for a period of decades, there is a divide between those practitioners interested in populations and those interested in treating individuals (indicated for example by Ferguson, 1999). Solomon describes the methods of clinicians interested in the treatment of individuals (GPs, patients themselves) as more likely to follow narratives prioritising early detection, which are more deeply engrained and have been widely believed for longer than the view that mammograms might cause more harm than good. Clinicians interested in the health of populations, by contrast, might point to evidence-based methods and to the results of some more recent studies, suggesting the

statistical benefit of regular screenings is such that the absolute risk reduction in early detection of cancers is 0.05 per cent.

Similar discussion about weighing the risks and benefits of interventions exists in the recent literature on behavioural insights, raising questions about the justification of carrying out interventions in terms of their risks to individuals (undermining patient autonomy, damaging psychological wellbeing) and benefits to the public (better patient outcomes in terms of group morbidity and mortality). In the case of regular mammography screenings, the absolute mortality risk reduction of 0.05 per cent is described as insufficient reason to influence people to attend screenings that are likely to have negative effects on their psychological wellbeing. Particularly when individuals were not independently motivated to attend (*i.e.*, they were motivated to attend by a nudge which suggested they might be at risk of death), the absolute risk reduction does not seem sufficient to justify persuasion that could increase individual psychological stress, for example about the fear of comparatively unlikely undetected cancers.¹⁹⁴

The concern about disproportionate risk of interventions, as in the mammography case where there appears to be the risk that screening might identify and cause the treatment of cancers which might never be otherwise noticed, exists also in arguments against the inclusion of enhancement in medicine. Identifying and (sometimes painfully) treating an illness which might never otherwise cause harm to the patient is a concern for increasingly accurate and predictive medicine, where treatment risks causing more harm and worry than patients would experience if undiagnosed conditions were never discovered. The narrative that early detection saves lives might be out of date, given the increasing accuracy and sensitivity of methods of detection, which have changed significantly since the 1960s and 70s when the narrative became well-established in the treatment of cancers, among other illnesses.¹⁹⁵ Narrative medicine is usually the most accessible to patients in a public health setting, where the context of habits, responsibilities, and group preferences do not easily allow for individual laypeople to scrutinise the data and arrive at their own scientific best analysis of risk. For this reason, primarily, it is in the interests of public health authorities to maintain clear narratives about the best course of action for *people in general* while allowing for exceptions based on individual differences in circumstance and preference.

Enhancement medicine, as the subject of ongoing discussion within the professional domain seems better addressed in terms common to public health engagement, to avoid erosion of the authority of the professional domain, while technological interventions towards

recognisable goals are still in earlier stages of development. As Hughes reasons,¹⁹⁶ it is important that transhumanism should engage with democratic processes: ‘only believable and effective state-based policies to prevent catastrophic consequences from new technologies will reassure skittish publics that they do not have to be banned’. It is not a contradiction for transhumanism as a libertarian project to rely on established political and social cohesion to cultivate better versions of its narratives. Transhumanist narratives about health and wellbeing should aim to match the progress of technological innovations, though sometimes public opinion does develop independent of the firm consensus or support of medical professionals about the uses of innovative technologies. In the absence of consensus about a line between public health and social interventions which might tend to improve public health over time, there appears to be a grey area. More conservative accounts of public health express concern that this grey area could be easily exploited by ill-intentioned governing bodies to enforce social change, for example with the implication that divergence from desirable behaviour constitutes a social ill which can be treated for the general good.

Though the habits and lifestyles of individuals, their health-span, and average lifespan may be strongly correlated, individuals with certain habits are not distributing a public good among others. Neither do campaigns to encourage good habits seem as invasive or inescapable as measures that mandate treatment (*e.g.*, compulsory vaccination) or preclude activities (*e.g.*, legislation that makes certain classes of drug illegal to possess). Intervening on societal causes of behaviour, which appear to fall into the grey area present cases it may be best to judge individually rather than against a rigid justification framework. Between, for example, the provision of clean drinking water or vaccine mandates, and measures intended to encourage people to drink more water or to engage with vaccine programmes voluntarily, there are norms and defaults guiding lifestyle and engagement with structures of healthcare which can be addressed less directly. The way individuals engage with medical knowledge and with changing best advice from experts also appears to affect trust in medical authority and engagement with public health. Solomon’s analysis of the ‘developing, untidy, pluralist methods’ involved in ever-changing medical knowledge guiding best advice suggests that the authority of public health judgements comes when the evaluation of risks and benefits are supported by narrative consensus, if not by formal consensus of all professional bodies, and are liable to change.

Additionally, public health interventions, even those concerning the protection and distribution of agreed public goods can often limit individual freedoms. Dworkin’s defense of what might otherwise be straightforwardly paternalistic interventions, that their benefits

can only come by group action and justify that action, also serves arguments in favour of human enhancement medicine as public health. Technological avenues for human enhancement medicine with the most global reach and likelihood of equitable distribution, those which aim at the most widely agreed upon improvements to the human condition, tend to aim at increased robustness in old age, and disease-resistance (increased health-span). There is no global consensus about these interventions, nor about their goals, however. Both the provision of clean drinking water and maintenance of healthy everyday habits (diet and exercise, smoking, or alcohol intake, for example) are predictors of lifespan and health-span, but there is less controversy about the need to standardise and protect the provision of clean drinking water. The WHO and other global health bodies agree that people ought to have clean drinking water and that this should be protected by law. The consequences for governing bodies of failing to supply clean drinking water to citizens are reinforced by law in many places, for example the Safe Drinking Water Act (USA, Public Law 93-523 since 1974), Water Act (UK, 1989, placing the provision of clean drinking water as a responsibility of the Secretary of State for the Environment). While certain habits may be shown to predict bodily conditions associated with high morbidity and mortality, associated morbidity from poor lifestyle choices tends to be more controversial in terms of legislation. Lifestyle as the cause of ill health tends to involve narratives about individual responsibility, guilt, and shame, in a way lack of access to generally agreed upon public goods does not.

For example, although higher levels of visceral and sub-dermal body-fat are often identified as increasing the risk of comorbidities and mortality, body weight is a controversial focus of public health interventions, and narratives defining a higher body fat percentage as a state of ill health independent of comorbidities remains controversial. Normative judgements about the responsibility of individuals with a higher body fat percentage for consequent illness vary culturally, involving different ideas about, *i.e.*, responsibility and blameworthiness for illness, personal health and wellbeing, and privacy. Because habits tend to be considered private rather than, for example, the direct result of policy failures or lack of access to generally agreed public goods, a more stringent account of public health as only the provision public goods might allow the exclusion of more difficult and grey areas whose controversies might risk undermining some of the justifications of public health authority.

However, given evidence from clinical studies in support of the effect of lifestyle and habit on health and lifespan, it is increasingly possible to predict and weigh the risks of reasonable public health concerns before their ill effects. The health of individuals does not exist in isolation from the health of others, as parents with bad habits seem more likely to have a

negative influence of the habits of their offspring, and the narratives of one generation seem likely to influence those of the next. In healthcare as in climate change and social care, apparently private matters contribute to the social norms that admit inter-personal responsibility, and do not concern only individuals. Some of the grey areas, no matter whether they are likely to be controversial, share other important characteristics of public health issues.

This section has addressed a difficulty in identifying transhumanism among public health problems open to special ethical justification for intervention: stronger definitions of public health which allow and justify social intervention also seem to exclude transhumanism from public health. Many problems requiring complex behavioural and social interventions are also public health problems, given causal relationships between lifestyle and health-span, social situation and lifespan. Social aspects of individuals' health and wellbeing are more often sensitive and controversial, particularly when governments intervene, and this adds complexity but does not change the status of the interventions as interventions for public health. For transhumanism, just practice of enhancement medicine relies even more than therapeutic medicine on authorities distributing resources evenly and protecting the autonomy of patients engaging with the unfamiliar.

Moving towards descriptive models and away from evaluative models of global disease burden.

Measuring the burden of disease morbidity and mortality is complicated by social and evaluative factors, particularly for the illness and death of individuals who are close to the population's life expectancy. Ageing populations in better living conditions tend to die from some causes at a higher rate than from others for complicated social as well as purely medical reasons. The rate of cancer, for example, tends to rise in ageing populations with better living conditions and better healthcare, and seem likely to include the deaths of individuals who might otherwise have died years earlier from other causes (for example in preventable industrial accidents or from treatable infections). If people live long enough, they become more likely to die as the result of age-related decline, which tends not to be treated as preventable in the same way. Causes of death are often divided and weighted by age groups¹⁹⁷ to produce a standardised weighting, to allow causes of death to be compared meaningfully between populations with different age distributions.

When a death is considered preventable, it follows that those years of healthy life which might have been lived are in some way *lost* as its result. Measures to prevent the loss of healthy life years or to add healthy life years justify some authoritarian intervention to identify the causes of loss and to disrupt and prevent predictable harm coming to individuals. A public health issue can be identified statistically, in terms of the burden of disease placed on a population, when all the years of healthy life lost to a certain cause are combined. QALY (Quality Adjusted Life Years), HALY (Health Adjusted Life Years), and DALY (Disability Adjusted Life Years) are evaluative models which have been used by epidemiologists and public health authorities for several decades, to identify areas of concern for public health, tracking the burden of disease and the risk of death from various causes. Models like QALY, HALY, and DALY have been used to support global disease control and prevention since the 1970s, and have been standard practice since the 1990s, replacing Crude Death Rates (CDRs¹⁹⁸), and age-specific CDRs which also tracked age-group risk, for example infant mortality rates. Although medical science is advancing the available treatment for cancer patients, and cancer is recognised as one of the largest global causes of death, older patients are more likely to have receive worse prognoses and more likely to receive palliative care than aggressive treatment.

QALY and DALY models aim to identify rates of preventable death, against demographic measures like life expectancy, to assess the burden of illness on populations. There has been a philosophical shift in epidemiology in recent years, with many seeking to move from evaluative models of death rates and health loss like QALY, toward more descriptive models. The calculation of lost life years involved in the DALY model is less evaluative and more descriptive than its predecessors but has also been criticised on the grounds that it perpetuates hidden assumptions about the value of the life years of certain individuals.

Though they give a more detailed view of the burden of disease, evaluative measures like QALY and DALY have been criticized for leaving resulting policy decisions open to the influence of bias about whose health is most important in a society. Because of underlying assumptions which are not necessarily explicit in the calculation of life years lost, models of disease burden in general risk valuing some individuals' health more than others (favouring young adults over young children or the elderly, for example) or ignoring the social and psychological dimensions of health life years (like meaningful work, mental health, social belonging). There is also the risk that models of health in terms of burden value a benefit felt immediately over a benefit that will be felt at a future time (including benefit to future generations) because of the cost involved. Finally, there is concern that the way burden is

weighed in statistical models might risk perpetuating generalisations about the contributions of those with disabilities, marking some conditions of health as less acceptable than others, and perhaps perpetuating hidden assumptions and misconceptions about living with disability.¹⁹⁹

Models of disease burden do not only record the relative health of populations and the most common causes of death in demographic groups, but also help determine which deaths should be tackled with the most urgency by public health measures. Organisations working with the results of global disease burden studies operate with constraints of time and funding, and it is necessary to assign priorities to the focus of public health interventions, when dealing responsibly with causes of preventable death.²⁰⁰ As discussed in the previous section of this chapter, there is a limit to the interventions by authority that democratic societies will tolerate, and not every cause of sickness or suffering can be called a public health issue and tackled accordingly, without other problems arising.²⁰¹

Priorities, in terms of public health, should²⁰² change and can change to address certain behaviour given new findings from medical science. This change can even happen quickly, given the urgency or immediacy of a risk. For example, from the mid-1950s when cigarette smoking was identified as a cause of avoidable deaths from lung cancer, cigarette smoking started to become the focus of public health efforts which aimed to stop heavy smoking and to prevent an epidemic of deaths that are preventable according to global disease burden studies. The risk from cigarette smoking is not only that individuals are more likely to die at a younger age than their expected lifespan, but also that their health is likely to decline more quickly from a younger age than non-smokers, likely culminating in an earlier death. The health-span of individuals who die from lung cancer is likely to be shorter than the health-span of the same individuals if they had not smoked, and the years of illness and suffering that could be predicted for smokers helped policymakers weigh the burden of cigarette smoking on public health. The weight of the burden helps determine how much time and money should be spent on prevention and social care. This is important when considering the value of enhancement and preventative interventions to change the health-spans of human beings, in part because the Adjusted Life Year models still consider age at the time of illness or death and evaluate the burden of suffering from disease accordingly. Public health interventions are those aiming to remove or reduce globally significant causes of lost life years and lost adjusted life years in poor health. Due to the scale of their impact on global public health, the most significant causes of global death and disease burden also tend to be problems that are immediately recognisable and unanimously undesirable, for example

cancer, heart disease, or malaria. Efforts towards the eradication of these burdens do not tend to cause controversy, regardless of variations in social norms.

In general, models of morbidity and mortality from different causes have weighed the number of deaths from certain causes against mitigating factors, such as the age of the individual at the time of their death. Eight core domains of health were adopted in the Global Disease Burden (2000) study, measures of health: mobility, self-care, pain and discomfort, cognition, interpersonal activities, vision, sleep and energy, and affect. It is more likely that elderly individuals will experience losses that constitute health loss in these domains, and this is widely accepted and expected by individuals as an inevitable feature of old age. Some argue that the age-weighting involved in Adjusted Life Year models of disease burden are unfair, because years of life are just as valuable when they are years lived by the elderly as by the young. Others suggest that the models are right to maintain age weighting on the grounds that it better reflects social expectations and priorities.²⁰³

The disagreement about whether ageing should change the evaluation of disease burden is part of a wider philosophical disagreement about how evaluative models of morbidity and mortality should aim to be, and to what extent evaluative assumptions (from social expectations and priorities, whether explicit or hidden), should drive public health priorities and justifications. Evaluations guide policy because people must decide what they value and prefer, and people have ideas about what they value in the behaviour of others. In public health, as in other policy decisions, the preferences and values of a society are usually involved, whether deliberately and explicitly or implicitly as bias. Arguments in favour of evaluative elements of disease burden and morbidity modelling state that factors like age do make a difference to how losses in the core domains of health are experienced by individuals. It is more distressing in some respect, this argument would suggest, to lose mobility as a young adult than it is to lose mobility in old age precisely because social expectations of mobility are reduced in old age where bodily decline is expected and normal. Older individuals are less likely to miss expected normal life experiences because of losses in the core domains of health, for example in work or reproduction. This argument considers the mental and emotional health of individuals, beyond the physical effects of losses in the core domains of health.

Further, arguments in favour of evaluations in models of disease burden and death rates suggest that biases exist regardless of whether they are 'built in' to the way the models produce weighted data, explicitly. Bias and hidden evaluations would simply continue to

effect policy directions, whether they were involved in the calculation of *i.e.*, DALYs, or not. Policymakers would be no less likely to continue to discount the experiences and diminish the suffering of certain demographic groups like the elderly based on common assumptions, or to continue to value future benefits much less than current ones. At least, supporters of a moderately evaluative approach to weighing burden suggest, in the case where the model involves making explicit assumptions underlying evaluations that guide policy, biases are transparent and acknowledged.

However, when weighing disease burden differently in old age and about burdens and benefits in the future, the explanation that society tends to agree with these judgements seems insufficient. Though it may be true that there is a socially normal expectation of losses in the core domains of health in old age, social agreement does not seem sufficiently good as a reason to weigh losses less heavily when they happen to older people. Although it is true that losses in the core domains of health are statistically expected in old age more than in youth, it is not the case that losses are expected or welcomed by individuals about their own ageing bodies in the same way, nor that every individual has the same threshold for describing old age where losses are to be expected.

Further, social norms about what individuals should expect to experience and enjoy in old age (work, travel, relationships, creativity, sport) are likely guided by the losses that were common among recent previous generations. Keeping ill-defined assumptions about what is normal at any age, as a part of the weighting of disease burden, risks becoming circular if the justification for weighing losses as less distressing or inconvenient is that older people do not wish or expect to partake in as many activities or experiences in old age. Social norms might make it more likely that individuals become stoic enough to expect and accept the losses of old age and to adapt their expectations about activities and potential quality of life accordingly. Life expectancy is a relatively accessible, 'everyday' metric which tends to be familiar to individuals with no expertise in global disease burden studies.²⁰⁴ Nevertheless, beyond individuals employed in actuarial sciences or pension planning, a small change in the expected lifespan of certain demographics, from bi-annually updated global data does not seem likely to elicit a change in general attitudes about what constitutes old age. There is a delay, almost by definition, for social norms to catch up with new data, and they should not be expected to determine when certain losses in the core domains of health should be accepted more or weighed as less burdensome.

There are concerns about the lack of sensitivity in QALY. Though DALY risks inaccuracies by inaccurately describing some losses as indifferent which should be considered burdensome, QALY risks generalising the experience of conditions which might affect individuals very differently. Losses in some of the core domains of health are evaluated as more sensitive, damaging, and burdensome, according to the public (layperson's) view. Assumed differences are likely also to vary according to other demographic differences like age, sex, and location. Burdens are difficult to assess in objective or abstract measures and weighing burdens by demographic risks generalising people's diverse lived experiences of disabilities. Evaluations and comparisons of disabilities which do not accurately represent the lived experience of individual people may inaccurately identify certain differences in ability as more or less burdensome in terms of public health than they realistically are.

Although many of the primary causes of global disease burden and death, such as cancers, heart disease, or malaria, are commonly recognised by laypeople, tend to be unanimously considered burdensome and non-valuable states which should be eradicated, if possible, others are more controversial. In the transhumanist primer, I described a grey area that exists for attempts to divide medical interventions between enhancement and treatment when the division is based on the treatment of disease or the normal function of the body.²⁰⁵ Genetic predisposition to illness and measures taken to reduce the likelihood of illness from genetic variations might appear to fit more closely the model of general medicine than do the accumulated causes of ageing and age-related decline on disease-based and normal function accounts of medicine, but the professional domain also guides attitudes to treatment. Public health initiatives dealing with behavioural and emotionally charged causes of disease burden should be more careful not to exacerbate guilt or shame associated with being unwell. Public health burdens are more ethically framed as group responsibilities than as individual failures or vices, therefore. For human enhancement, the concern is that if, as described above, genetic predisposition to illness were seen to share the characteristics of other kinds of public health problem (with similar kinds of generally agreed public narratives about health), it might result in guilt, shame, pressure to conform genetically (as is experienced with disease burden associated with body fat, smoking, alcohol intake, sedentary lifestyles, etc.).

Theories of libertarian paternalism often aim to redirect people to behave in ways that better match their preferences and values, to overcome implicit biases and hidden assumptions that would otherwise be confounding factors. The approach has been criticised for the hypocrisy involved in supposing that experts in human behaviour have authority to redirect

people away from harmful and unproductive biases in their unconscious decisions, despite being vulnerable to the same biases. This criticism of nudges²⁰⁶ addresses a concern about paternalist interventions more broadly, that those in authority are not free of the biases they aim to prevent in others when they decide on policies and interventions.

The discussion above, between evaluative and descriptive models of disease burden demonstrates a promising way forward for this challenge to nudges in public health. The move to avoid some of the more forceful biases, against older people and far off future times, that have been driving the distribution of public health resources involves a more holistic narrative account of life which rejects the atomism of QALY. This approach from Dworkin defends the seriousness of a *good* death, as one with dignity which is fitting given the value and quality of a life.²⁰⁷ The narrative approach from Dworkin emphasises that suffering does not become any less worth avoiding after more years of life have passed, as the narrative of the individual life becomes fuller in old age.

Dworkin's holistic narrative account, in opposition to models like QALY which do not prioritise palliative care for terminal patients, for example, also values autonomy and freedom. Palliative care prioritises dignity and freedom from pain towards the end of life for individuals who also rely on the care of other people. QALY and DALY models are less sensitive to the lives and experiences of individuals, but offer a clearer framework for allocating the financial resources needed to address public health issues. Miriam Solomon's developing untidy methodological pluralism of medical knowledge²⁰⁸ would better maintain the balance between liberty and protection from harm that is involved in treating diverse groups equally, given the need for both atomistic and holistic measures of value and burden at different points of delivery of public health.

Modelling deaths within a population, by cause, age at death, and by the potential for prevention, as more than crude numbers (crude death rates or CDR²⁰⁹), contributes to the direction of public health. Identifying areas of public health that are most in need of attention via their social and environmental causes, as well as by identifying the bacterium or virus at fault, models of morbidity and mortality help policymakers decide how urgent a problem is, whether public engagement would likely aid in its reduction, and how many resources should be set aside to improve the average risk of death. The evaluative aspect of disease prevention has been important because they allow greater insight into the need for action if some deaths are 'postponed' rather than prevented, or some deaths might be considered worse than others, which rely on assumptions that are not easily reducible to even age-adjusted

data. For example, morbidity from an illness in children under five years of age might be measured in a simple and transparent way and compared with morbidity from the same illness of people in older age-brackets. This comparison of rates of morbidity in older and younger age-brackets might help identify a discrepancy of risk, so that public health interventions could be used to target and assist those most at risk, however the data do not consider the years of life lost to children who die before the age of five. The introduction of evaluative measures to modelling aimed to harness assumptions that also inform policy decisions, for example that the death and loss of years of life of children under five is worse than the death of older individuals with fewer remaining years to lose.²¹⁰

Transhumanism appears to cause a problem for this method of identifying public health problems as death in old age (sometimes discussed as death *because* of old age), is not currently considered preventable death. Many individual causes of death in old age, including those resulting from the process of ageing, are already considered preventable, however. For example, several cancers become more likely to develop in old age as cells become less able to replicate without problems, and mutations become more likely. The rate of cancer morbidity tends to increase with age, regardless of other factors like lifestyle, and risk of cancer mortality also increases in older patients.²¹¹ Because of the prevalence and effectiveness of medical interventions, the accepted definition of a death in old age, synonymous with a death that was not preventable, has changed over time. So, the integration of transhumanist assumptions into frameworks for identifying important public health interventions does not represent a radical change in a fixed attitude to dying but a part of a continual shift, appropriate given changing medical technology.

In this section, I have discussed a challenge that exists for public health epidemiologists whose work guides the weight and direction of public-facing campaigns and projects. Evaluations and assumptions about death and suffering are involved in judgements about quality of life and medical best practice, given current models of disease burden. Assumptions about morbidity and mortality are not always fair or just, and rest on more general social biases about the quality of life available to individuals in different demographics. The professional domain has not always agreed about what counts as a disease or which are the most serious social burdens caused by disease, and it is not clear that scientific knowledge avoids the influence of cultural bias, even if evaluative measures are removed from epidemiological records of death. In the next section, I discuss a further and relevant challenge for cases when social norms do not agree with the direction of public

health campaigns: it can be difficult to determine how much pressure authorities should exert on citizens who disagree about what is best.

Maintaining libertarian paternalist balance with nudges: why not educate the public about their enhancement options?

Nudges are a good tool for public health policies, because they promise cost-effectiveness through their 'light-touch' interventions into the lifestyle choices made by individuals.²¹² Nudges are also useful for difficult public health cases where lifestyle choices are determined in part by narratives about what is valuable because of the way nudges can engage people's emotions to re-evaluate behaviour which had been the status quo. These uses of emotion and narratives, for example about commitments and the expectations of others in public health nudging, are not a clumsy or heavy-handed consequence of attempting to inform people, rather it is essential for ethical public health measures to remain effective that emotion and narrative are involved. In chapter A3, nudges were defined as follows

"A nudges B when A makes it more likely that B will ϕ , primarily by triggering B's shallow cognitive processes, while A's influence preserves B's choice-set and is substantially noncontrolling (i.e., preserves B's freedom of choice)."²¹³

Nudges occur as one method used in behavioural economics (also called behavioural insights theory) to influence the behaviour of people. Nudges in policy tend to aim to help people to do things they feel they ought to do, with as little restriction (for example legislation) as possible. Behavioural insights work can engage people's decision-making processes, conscious and subconscious in several ways, targeting several kinds of bias that occur in decision-making processes: messenger, incentives, norms, defaults, salience, priming, affect, commitments, and ego.²¹⁴

Behavioural insights policy work accepts the premise that acting neutrally is not possible, and so it is often unethical for policy to ignore the influence of small individual choices and stated values on other people and on progress in general. Behavioural insights work is not just about governments influencing individuals, therefore, rather it is about individuals influencing other individuals, with snowballing effects. It is sometimes more beneficial for authorities to allow these effects to continue unimpeded, for example where social values change independent of legislation (perhaps even culminating in petitions and protests to change legislation, once it no longer matches social norms). However, in more complicated

and less predictable cases, for example in the face of emerging human enhancement medicine, there can be no neutral position taken by authorities, and inertia seems worse than proactivity for health outcomes.

An example from the first report from the Behavioural Insights Team under the UK Institute for Government suggested that although automatic processes of planning and decision making serve adequately, most of the time, for making the daily commute to work, the same approach would fail to serve when making an unfamiliar journey.²¹⁵ When people make decisions or plan for futures that are not routine or familiar, behavioural insights interventions can choose either to prompt people to think in more reflective ways,²¹⁶ or else can re-arrange the choice architecture (analogously, the landscape in which automatic journeys take place) so that people are more likely to safely arrive at their destination using automatic processes alone.

A tradition of political philosophy²¹⁷ highlights the need for broad emotional consensus between people in a society about what is valuable and worth doing, for there to be justice. Establishing principles, writing them into law and upholding them with fairness by law is one part of the political project of maintaining just societies. The other part, defended most recently by Nussbaum,²¹⁸ concerns the maintenance of society through a culture of sentiment in support of the core principles of justice. In short, Nussbaum argues, it is important for people to care about things and want things to happen, in order for things to keep happening, and to avoid tension between the rule of law and the freedom of individuals to act according to their values.

Bridging the narrower emotional experiences of individuals and broader sentiments about public goods and public projects of cooperation, for example public health, and a continual critical dialogue between political emotions and political principles is necessary to maintain the balance, says Nussbaum²¹⁹ and in the absence of these measures, the balance between the demands of authority and the will of the people will not last. To relate Nussbaum's ideas to the problem for transhumanist public health in terms of a previous challenge for society in the face of technological revolution, I consider the case of documented widespread computer anxiety in the 1970s. I suggest the case demonstrates what can happen quickly and with lasting effect when people's emotions do not align with what is required of them by society, or their preferences do not seem to align with the demands made on their capabilities by authority, given fast-changing technology.

In the mid-1990s, meta-analysis of global studies into the causes and consequences of computer anxiety and technophobia undertaken by Weil and Rosen suggested a correlation²²⁰ between the 'comprehensive and well-integrated early computer education' and the proficiency and confidence of educators in computing, with the competence of students at university level, and a negative correlation of education and confidence of educators with the incidence of technophobia and computer anxiety.²²¹ Earlier, around the mid-1980s, years after computer anxiety, technostress and technophobia had been identified following the introduction of personal computers into office workplaces, The Council of Europe's Standing Conference of European Ministers of Education met to discuss the challenges of the technological revolution in computing and the internet. The 1989 Council were:

'DETERMINED that education should fully prepare young people to meet the challenges inherent in this emerging information society and to play their essential democratic role in helping to shape it;

CONCERNED in particular to help all members of society to have better access, through education, to the knowledge and skills they need in order to participate in the information society, and so overcome the effects of differences due to poverty, class, physical disability, age, gender, national, cultural or ethnic background or geography;

*CONVINCED of the continuing great importance of the historic role of education in building active independence of judgement and humane values.'*²²²

There are many parallels between the position of the education councillors, academic researchers, and the lay public in the 1980s in Europe and globally, whose quick response was essential to the successful and ethical distribution of computing and internet technology in the years to follow. The period, which was described contemporaneously as 'a technological revolution'²²³ consisted of the previous 30 years, with an acknowledgement that the pace of change was likely to continue its acceleration. The meta-analysis of global reactions to the fast pace of change between the 1960s and 1990s considered the effect of computer anxiety and technophobia, as the emotional consequences of a feeling of unpreparedness for emerging technology at the point of use, for individual laypeople. Present use of computers created 'self-critical internal dialogues' in university level students, globally. Contemplation of future use of computers were also found to effect individuals negatively in those experiencing computer anxiety. Technophobia or 'computerphobia'²²⁴ could even result in individuals feeling uncomfortable speaking about computers, and

individuals reported feeling unwilling to even think about computers or their use. These results were gathered at a time when the use of computers on university campuses was already becoming more commonplace, replacing earlier analogue library and archiving systems, with the requirement that students acquire a personal computer for study becoming more normal at university level.

Although education policy and the necessary changes to education, from a student as knowledge-gatherer paradigm to a critical thinking and knowledge-filtering paradigm, which is more familiar in modern approaches to education,²²⁵ were discussed by policymakers in the decade before the results of the meta-analysis were collected, educational approaches to technological, fast-paced paradigm shifts have a time-delay before they affect university age individuals. Further, the results from the 80s and 90s emphasise the role of the competence and confidence of educators as a factor determining how comfortable and competent students of computing would be likely to feel. In the 80s, some studies estimated that around a third of people experience major aversive reactions to computer use or technology.²²⁶ A recent review of the incidence of computer anxiety and technostress by Agogo and Hess²²⁷ suggests that neither phenomenon has dissipated entirely, even in countries where availability and familiarity of the technology are all-pervading. It appears that early computer anxiety and technophobia is more lasting and persists in the attitudes of many frequent users beyond mere apprehension about the unfamiliar before it becomes familiar. The persistence of computer anxiety and technostress affects the experience of many people, particularly given the continued inclusion of computing in more universal areas of human activity such as education, communication, shopping, travel, and medicine.

Of the three practical implications suggested by Agogo and Hess, at the individual level, individual traits which make technostress and computer anxiety more likely or more severe could be assessed and addressed more sensitively.²²⁸ Technostress and computer anxiety could also be addressed at organisational level (routine assessments and feedback to managers about negative impact of technostress) and at technological level (better interface design).

In the case of computer anxiety and technostress, negative effects for affected individuals and their employers, for example relating to employee retention and performance,²²⁹ have now been familiar and recognised as individual variations in employee competence for several decades. Efforts to address the negative consequences of insufficiently good outcomes for individuals who must interact professionally with now all-but-inescapable

technology might not be high priority for private organisations at the point where change requires additional expense or redistribution of resources. Assessment of individuals for traits which might put them more at risk for computer aversion could become a double-edged sword, putting the emphasis on employees being unfit for roles, rather than on employers as responsible providers of comfortable and productive places of business. Interface changes seem promising but would require investment by organisations not necessarily invested in the success of individual employees at companies which use the interface.

Education and further training are also expensive, time-consuming, and can come too late for adult individuals who are already in regular interaction with the causes of their technostress, or whose career trajectory has been altered by the requirement to engage with fast-changing technology. In the workplace, technostress can cause performance and mental health issues that are serious and undermine self-image, as well as blood-pressure increases and other physical health conditions with chronic exposure.²³⁰ It does not require a serious imaginative leap to suppose that a similar trajectory taken by the introduction of medical technology to the everyday experience of individuals would be devastating in similar ways. The challenge of increasingly technological medical interactions has already been identified.²³¹

Education is important, contributes to the ethical treatment of the public by authorities in times of technological revolution, and is beneficial to future generations. However, behavioural insights work suggests that education should be augmented with the 'goal of fostering the involvement of the private sector in public health promotion' where previously organisations with influence might have been wary about involvement in the preferences or decisions of individuals in their sphere of influence.²³²

Further problems raised in this chapter:

Transhumanism is not merely concurrent with general medicine in the way other technological successes like artificial intelligence, communication, or exploration advance the human condition alongside general medicine. Transhumanism, because it shared the principles of bodily wellbeing and longevity that have been aims of medicine for some time would more likely fail to be ethical as a separate and competing medical discipline. Accepting the new public health paradigm under which enhancements are a 'live' and generally

accepted option for individuals, could be assisted by use of nudges and expert advice targeted at areas identified by global burden studies. In the thesis introduction, I described the introduction of just such an approach to enhancement in the language of the paradigm shift literature as a Gestalt switch that changes the phenomenal world, though gradually.

- a) Workplace promotion of public health responsibility is not yet supported by ethical frameworks. This account of public health and public responsibilities to group health is too vague and too broad, it is open to criticism based on questions of autonomy and voluntariness, discrimination and distributive justice, privacy, responsibility.
- b) Donor organs and blood seem to be treated as public goods by efforts to 'collect and distribute' them efficiently. The more they are treated like public goods, akin to access to drinking water, the better the public health framework seems to fit them. Narratives about helping others by contributing tangible resources like blood or funding would not matter cases where the intervention was less clearly about distribution.
- c) Paternalism and propaganda remain too controversial in matters like enhancement, leaving them even more open to rash value judgements than existing frameworks are open to biases that diminish the suffering of age-related illness.

These frameworks and strategies are as applicable to the enhancement of human lifespan and health-span, which transhumanism calls the pursuit of negligible senescence, as they are helpful for established public health measures which tackle behavioural and lifestyle choices, and to established public health interventions which aim to encourage the uptake of vaccination and organ donation.

This chapter has aimed to uncover pieces of the behavioural insights framework that would be appropriate and supportive of transhumanist public health endeavours. There are tensions common to transhumanist ethics and other public health interventions which become politicised and emotive in the public eye. In the case of organ donation, the

transformation of donor organs and blood packs from personal to public goods was made gradually, using emotions like solidarity, pride, respect, etc. to bolster public feeling and change the structures of distribution and acquisition, over several decades (and more recently, with targeted 'quick' campaigning leading up to the legislative change to 'soft' opt-out).²³³

In response to the problem that not all accounts of public health are inclusive of social problems or the medicine of paradigm shifts, I suggested human enhancement had more in common with preventative measures (vaccines) and decisions about responsibility to others (organ donation) than to immediate medical interventions. These, more social public health measures require more robust ethical justifications of the target groups and the methods used, though even immediate medical interventions and the public health policies about their use have implications for others and for future people. In tackling infectious diseases during an emergency, public health strategies use concepts which are also available and familiar as generic public health concepts: solidarity, equity, trust, autonomy, equal moral respect, and vulnerability.²³⁴ Concepts like these need to be elaborated upon to be useful in specified crises and in the face of logistical problems. However, because they can be elaborated upon in accessible ways, and are already at the heart of public health discussion by the top level in planning research, development, and acute strategy, the discussion of public health as more than public goods seems more accurate.

In response to the problem that human enhancement moves the evaluative 'goalposts' that make ethical frameworks for judgements about the value of medical interventions possible, I point to the increasingly descriptive work of statisticians, recording morbidity in terms of DALYs, to avoid hidden assumptions including age. The shift in attitudes to illness, suffering and death in the DALY model has much in common with the Humanity+ project, and with the goals of transhumanism, in the attempt to separate evaluations rooted in social norms about ageing and acceptable suffering from models that report the major causes of death and suffering, globally. Measures like QALY and DALY have been based on assumptions, for example about the value of years of life during old age, which are not easily defensible, yet persist. Measures of disease burden and the biggest causes of preventable human death still rely on assumptions to some extent but are attempting to move towards describing rather than evaluating the causes of human death. This aligns more closely with the values of progressive models of human wellbeing, for example the transhumanist modified Humanist principles laid out by Humanity+.²³⁵

The remaining chapters of the thesis will continue to rely on the work of behavioural insights theorists to describe and evaluate the needs of the transhumanist project, within the ethical framework this section has aimed to establish from three developing strategies of approach to commonplace public health challenges. In the next chapter, I will move to defend a particular use of nudges in public health matters that require group participation and approval to be fair and successful. Where there is uncertainty about risk, as in transhumanism, other approaches to public health intervention fall short because they are too decisive for the available clinical evidence. Not all nudge approaches from libertarian paternalism are appropriate for emerging technology or controversial applications of medical technology and some might risk causing further alarm by engaging public opinion too soon, and without clarity.

Conclusion

This chapter has aimed to show that existing frameworks for bioethics, which already justify the careful use of authority needed to have public health, are available to advocates for human enhancement medicine where its aims are negligible senescence, extended longevity and extended health-spans for the global human population.

I resisted the concern that public health interventions, even if they need not be limited to the provision of publicly agreed goods like clean water, should aim to restrict individual freedoms in times of emergency or unusually severe disease burden. This concern presented a challenge to transhumanism and the implicit paternalism of increasingly technological medicine, used as default, which would limit the ethical applications of transhumanism as public health. Transhumanists resist the concern on the following grounds:

- i) Changes that would come about as the result of transhumanism, as heritable enhancement of human beings by means of medical technology, reduce and prevent globally significant causes of suffering. If the same suffering came about as the result of a toxic water supply, it would be considered a public health emergency, because of the scale and severity of illness and death.
- ii) This is a status quo effect, which it would be reasonable to adjust against in the face of faster change, using nudges, now technological advances in negligible senescence are predicted to change in a short timeframe. Human ageing and

senescence towards death carry a status quo bias which means resistance in the way people respond to progress, and which would not usually be tolerated if other causes of suffering could be predicted and prevented.

I began to address the concern that nudges to adjust against status quo bias rely on manipulations which do not respect the ability of individuals to reason about their situation. Nudges are also based on hidden assumptions, and therefore do not allow people to engage with their lifespans and health-spans with the kind of patient autonomy that is required of patient-doctor interactions. Narratives of individual health and healthcare are valuable not only for individual healthcare (including concerns about palliative care and patient consent) but are also important in public health frameworks. QALY and DALY, as methods of identifying public health problems risk either generalising the lived experience of groups (like those with disabilities), or unreasonably separating and discounting the suffering of groups (like the elderly or the very young) in favour of initiatives that would help already privileged groups. Assumptions about natural ageing and old age tend to be strong and well-established. The QALY and DALY literature highlights how many assumptions about what counts as a preventable death are involved in judging risk and the value of treatment. Individuals are unlikely to be able to exclude biases about illness in old age while these narratives persist, or to make fair decisions about what counts as a premature loss of life, when they make choices that will change their own health-span and lifespan at an indeterminate future point. A further concern about paternalism of this kind²³⁶ is that nudge theorists situate themselves as above such biases and heuristic mistakes. This is not unreasonable and, in the case of QALY and DALY, the influence of assumptions on expert judgements is clear and very well-documented. Though it is possible for policy makers to try to adopt a legitimate birds-eye view when guiding people through public health issues related to ageing that they do not tend to navigate individually without succumbing to predictable biases.

For example, people might fail to worry proportionately about mental decline in old age because memory loss and confusion are associated with old age as normal and acceptable parts of that condition. For those who would support public health interventions to support shared values, encouraging behaviour that promotes negligible senescence is a way to shift the choice architecture to help individuals frame medical decision-making in cooperation with other individuals. This is one reason nudges may be better placed to help with medical

and health decision-making than they are in less emotionally charged uncertain decision-making situations. Considering the effect on others and about the judgement of others are a part of making choices about heritable and species-evolutionary medical interventions, and it is appropriate that they should be framed that way, though not with a heavy-hand, at the point of individual engagement.

- iii) Transhumanist medicine, because of its comparative novelty and because of the varying accounts of what the future of enhancement medicine will be like, is at risk of being stifled by uncertainty and complexity, if narratives about enhancement are left unmanaged.

The paternalism of medical technology, in the sense that authorities know what is best on the basis of their superior technological expertise, relies on more novel and more specialized expertise than other medical interventions with similar public engagement challenges. Vaccines, where the medical intervention and the social idea are more familiar to the lay individual, have more accessible narratives attached, for better or worse, which have accumulated over time since their development. The public may be likely to support the medical treatment of ageing once they feel strongly that such things are not inevitable human experiences, or valuable to human experience; and conversely people are more likely to feel strongly that such things should not be inevitable widely available and more widely supported medical treatments are the norm. In both the older literature on paradigm shift and in the behavioural economics literature, this apparently circular progress can be prompted or catalysed. A nudge as straightforward as a news report on the results of a clinical trial might highlight the positive outcomes of medical research into human enhancement, but without more sustained efforts to engage public opinion, new evidence does not carry much narrative weight.

I described the importance of emotional and supportive nudges, as overtures to more educational interventions on behaviour with reference to the way the introduction of workplace computers led to a sharp rise in computer anxiety in the late 20th Century. To understand why factual reporting on changes in emerging technology, for example the publication of scientific evidence as studies collect their findings, is insufficient to nudge public narratives about enhancement, I compared the case of rising computer anxiety in the workplace with the challenges facing the implementation of a transhumanist public health.

Patients, like office workers, are likely to be at an epistemic disadvantage at the point where engagement with the technology is required for their normal life to continue. At the point of first interaction with emerging technology, exposure to factual reports about efficacy or replicability, even accompanied by the publication of the outcome of consensus conferences, are unlikely to suffice. Unmediated apprehension about the pace of global change is likely to influence whether individuals are cooperative or uncooperative, consenting or non-consenting, amenable to change or more hostile, in medicine as in computing:

- iv) In workplaces and in medical care, nudges lend themselves much more naturally to broad effects that do not discriminate unfairly against already vulnerable groups. Restrictions and penalties can single out undesirable behaviour and become divisive.

The computer anxiety case shows a real-world situation where power imbalance led individuals to feel they had little choice but to comply, and this caused discomfort, anxiety and stress about change driven by technology. Because of practical social pressures related to training, income and self-development, the introduction of computers in the workplace was not smooth and has since been the subject of study as the 'wrong way' to introduce technology to populations. This identifies a problem for an account that attempts to talk about public health nudges in the way that 'staircase' nudges are often conceived. Further, financial incentives, though they can be very effective, put more pressure on those with lower incomes. Incentives are generally more effective than penalties, however, and the kind of pressure incentives tend to apply to vulnerable groups is less harmful than similarly weighted penalties targeting vulnerable groups.²³⁷

- v) Behavioural insights work is not completely predictable and its results sometimes have unforeseen side-effects. So, there is a risk that nudges backfire or could have no effect, and it is harder to account for individual differences and to present emotive nudges sensitively in settings with larger target populations. Some individuals may be 'pushed' or unaffected by an intervention which is an appropriate nudge for others.

Individual everyday decisions that affect health cannot all be monitored or managed by law or by nudge, and the design of public spaces is more complex than in spaces where the user group is smaller and more predictable. However, given narratives about outcomes, preferences, tendencies and normal behaviour, even without true consensus, coordinated social responsibility a real possibility. Individuals guided by generally accepted narratives about the right thing to do might have less freedom to avoid the nudge space altogether, for example medical practices, workplaces, or schools, but are less likely to be unfairly restricted than in cases where decisions made privately are incentivised or prohibited by law.

In this chapter I argued that the challenges described in iv) are the more serious ethical concern for nudges that aim to reassure and to promote positive action. Neither iv) nor v) seem to pose a problem *exclusively* for human enhancement as public health, as grey areas already exist. True consensus is not always possible, but there can be narratives that are accepted as the best fit, socially. Even when it is known that not all people are similarly knowledgeable on the best course of action open to them, strong narratives about the best values and choices make public health participation more accessible to more people. In the organ donation case, particular groups were sometimes targeted to encourage and influence their feelings but targeting and care taken over demographic differences in pursuit of positive healthcare outcomes does not amount to Propaganda by most widely accepted definitions in the literature.²³⁸

In the next chapter, I will expand on the claim made in this chapter, that nudges can be used in public health settings to offset generalised issues that are not necessarily urgent for the individuals who are nudged. Pre-arranged screenings, which can reduce the relative risk of serious illness for individuals who receive the pre-arrangement letters, are targeted nudges. In the next chapter, I expand the claim the breadth and generality of the goods at stake in the organ donation case mean it can often be better for public health authorities to justify positive nudges, even to promote emerging or less familiar conceptions of the public good, than for authorities to attempt to restrict behaviour. Regarding the gradual change, from organ donation being considered optional, personal, and quite exceptional, towards organ donation as a better understood process affecting societies but expected of individuals as default, I suggest that nudges can also be used to change which public health contributions fall into the category of contributions to national resources, and so change which stakes are considered tacitly to be acceptable.

The next chapter will build on the argument of this chapter that the imbalances of authority and autonomy involved in healthcare nudging, affecting human self-evolution, are justifiable in the face of uncertainty. This is particularly just given the persistence of outdated layman's assumptions about ageing which create bias towards conservatism. I argue that a similar framework to the behavioural insights-led series of nudges that recently have been successful for organ donation in the UK, could easily be adjusted to address foreseeable problems facing human enhancement medicine, as it approaches human trials under increasing global scrutiny. These interventions would not require the public first to accept or understand a uniquely transhumanist ethical framework for medical care, and this would not prevent the right kind of informed consenting participation.

Organ Donation as a Key Example of Behavioural Intervention

Why pursue nudges before approaches based on education or legislation, if democratic egalitarian transhumanism is the goal?

Introduction

In chapter 3 I asked, ‘What are the social norms and expectations of health that make bodily enhancement and longevity unappealing, despite relevant improvements in medical science? Are there good reasons for maintaining any of these norms and expectations about health, even if they slowed the transition to a more responsible public health paradigm?’

Nudges can make rapidly changing medical technology more socially responsible and less risky. Good uses of nudges can help people improve their risk-literacy when faced with change and uncertainty in medical science.

Hughes²³⁹ identifies the need for transhumanism to address negative attitudes by the inclusion of genetic enhancement in general medicine: ‘Overcoming popular resistance to technology will require not only assuring publics that they are safe and will not be forced on anyone, but also that there will be universal, equitable access to their benefits through public financing. In other words, genetic choice and enhancement technologies must be included in a national health insurance program.’ I agree with Hughes but suggest overcoming resistance to technology involves more than the promise that enhancement will be integrated into existing healthcare systems.

In this chapter I defend the use of nudges to overcome one ethical challenge for public health interventions already within existing healthcare systems that arises when the public are uncertain about the best way to arrive at a generally agreed good. When people are ambivalent about the best health outcome at the time important decisions are made, public health measures can guide preferences and make sure preferences result in better decisions, and they can also help solidify the normal way to achieve a public good.

Critics of transhumanism as human enhancement ask to what extent it is reasonable to persuade people to enhance their own bodies, for example to periodically undo the effects of ageing, particularly if people are uncertain or ambivalent about doing so. In this chapter I aim to reframe the question ‘what makes enhancement unappealing, despite improvements

in medical science?', asking instead to what extent it is reasonable to expect people to engage with changes in medical science that aim to alleviate their future suffering and the future suffering of their descendants. I argue it is reasonable to expect people to engage with negligible senescence to the extent that it is predicted to become the best way to arrive at many generally desired public goods. I refer to existing public health policy for examples of how public health encouragement about something uncertain can be carried out sensitively and make a significant difference to public engagement without undermining autonomous democratic participation.

It is not unusual in public health for a public health intervention to focus on discussion of the broader social aims of a society that wants to use certain medical technologies and, I suggest, it is not ethically dubious for public health policy to avoid drawing attention to technological interventions on the body. Interventions on the human body and interactions between body and technology tend to be sources of anxiety and fear, in healthcare as in other areas of life, and highlighting bodily fragility alongside technological interventions seems likely to cause unnecessary unease. I aim to show that there is sufficient well-established justification for an approach to public engagement that avoids discussing bodily interventions, as evidenced by past public health measures. Vaccination programmes, organ donation and blood drives, social support of reproductive medicine, and support for nationalised healthcare more generally tend not to highlight the procedures that are carried out on bodies by technology, and it is no more socially irresponsible for enhancement to be presented this way. Narratives about solidarity, trust and equity are important in public health to help autonomous individuals make better assessments of risk. Public health problems, like managing the introduction of enhancement into medicine, can often be better addressed by good behavioural insights work than by either programmes of public re-education about the future of medicine or legislation to enforce public health responsibilities relating to novel medical technologies.

Background

Some projects can only be undertaken successfully by coordinated groups of individuals working together. In the thesis introduction, I suggested that Transhumanism must be cooperative and species-wide or would fail to meet its objective of improving the core capacities of human beings, instead creating a greater divide between enhanced and unenhanced ('haves and have-nots'). This is a concern that has been expressed by critics of

transhumanism, and which poses a serious ethical challenge for transhumanists who suggest that enhancement is not merely a project for the few interested tinkerers to experiment on their own bodies,²⁴⁰ but has real potential to reduce the suffering and improve the wellbeing of the species. For transhumanism to achieve controlled and deliberate species self-evolution²⁴¹ it must be inclusive, democratic, cooperative, and global.

Larger cooperative projects often rely on more comprehensive organisation and enforcement strategies than do projects that rely on few individuals acting together. As Dworkin suggests in his defence of the use of authority in some public health endeavours, the problem of free riders and individuals who benefit more if they do not behave in the group interest means that some goals can only be achieved if there are restrictions of autonomy to ensure that enough individuals act together.²⁴² For example, the intervention of authorities to incentivise vaccination is important and justifiable to reduce the incidence of free riders so the population can achieve herd immunity. If individual vaccination were enough to ensure the safety of all individuals who acted cooperatively, the justification for intervention would be reduced and would be a more straightforwardly paternalist attempt to make sure individuals choose the option best for their own wellbeing. Paternalistic interventions of the latter kind are not unheard of, one example could be compulsory seatbelts in private vehicles, but are unusual in medical care where patient autonomy tends to take priority. Risky or harmful actions that only risk harm to the individual making the choice, for example smoking in private residences, alcohol and drug consumption,²⁴³ suicide and self-harm²⁴⁴ are not forbidden by law in most democratic societies, though they tend to be subject to other restrictions (for example standard unit measurements and alcohol sales licences), taxes, and close monitoring. Restriction by law tends to aim to prevent harm and, in democracies, there tends to be an assumption that lawmakers should be careful not to restrict individual freedom without good reason from predicted harm. Dworkin's argument, that some public health projects should use of authority to restrict freedom and promote a group project with goals that are only attainable by cooperation, also operates on the assumption that the restriction of freedom should aim to prevent harm, albeit in a wider context: some public health measures prevent harm to individuals, while others prevent harm to societies.

Transhumanism as medical human enhancement will likely change the core capacities of future human generations. The uses of human enhancement to tackle global disease burden and reduce suffering in old age is a matter for coordinated national and international intervention and should be treated as a public health paradigm, as argued in B1. Many other

public health projects do not address illness in individuals at the point of illness, but instead prioritise prevention, with governments improving conditions for future generations, and individuals acting on the understanding of risk as a product of statistical likelihood within their population. The social, cultural, and ethical interventions made by existing public health measures also have multi-generational effects and sometimes the multi-generational effects of measures taken are uncertain, given the potential of medical technology to change and improve with time.

I will suggest it is not unusual in public health, nor is it especially manipulative for a public health intervention to focus on discussion of the broader social aims of a society that would want to use enhancement medicine, not on the bodily interventions involved in treatment or the novel technology that would allow them to be carried out. I aim to show that there is sufficient and well-founded precedent for this approach in existing public health measures on issues like vaccination (herd immunity), organ donation (preventable deaths on waiting lists), reproductive medicine (the difference between availability and accessibility), and support for nationalised healthcare more generally. Narratives about solidarity, trust and equity are important in these cases and nudges are well placed to support and promote individual autonomy by guiding the public through a complex and unpredictable shift in medical technology.

The shortage of viable organs for transplant has been a factor underlying a preventable disease burden in the United Kingdom. A shortage of available organs for transplant has been shown to contribute to the death of many individuals on the waiting list for organ transplants each year. The adverse effect of the organ shortage on quality of life: health-span and lifespan, was identified first in 2006. A project to tackle the shortage of organs for transplant, and infrastructure problems that were undermining effective and timely intervention to acquire organs, gained government support and public health status between 2006 and 2008. This meant, among other things, more funding for a designated Organ and Transplant service operating nationally, and public-facing campaigns to draw public attention to the issue, to encourage voluntary registration of donors during life, to reduce the number of families who would act to prevent donations taking place by overriding registered donor status following the death of a loved one. Like vaccination or healthy eating programmes, organ donation benefits the social group more than the donor individual and can strengthen social ties, but unlike vaccines or healthy eating, organ donation does not tend to benefit the donor individual at all, and benefits a comparative minority of actual patients each year, proportionate to the whole population. Similar to other invasive medical procedures that

involve the transfer or removal of human tissue or its use beyond the biological processes of the body without technological intervention, the idea of taking organs even after death has historically been a source of fear and sometimes disgust.²⁴⁵

The implementation of new laws to govern individual choices relating to public health might be effective regardless of the preferences of the public that are expressed in public opinion polls. Assuming that most people are more likely to abide by the law than to engage in unlawful behaviour knowingly, legislative change can be effective as a change to the default social expectations placed on people, and sometimes the public morality shifts in response to the law, rather than the reverse being the case. However, the overuse of legislative authority can quickly undermine trust and it is not clear that most people do abide by laws with which they openly disagree, at least not indefinitely, without protest, boycott, or social unrest. Given excessive implementation of laws driven by medical science, loss of trust in the beneficence of authority could also create dangerous loss of trust in scientific best practice, and cause rifts between perceived normal and expected choices and the scientific medical understanding of good choices. In general, more cautious measures to increase engagement with change in medical science seem more likely to be successful than sudden or forceful measures, and measures supported by generally accepted social values seem more likely to last.

In the case of emerging medical technologies, for example heritable gene editing of both adults and pre-implantation embryos, I suggest it is reasonable and ethical to discuss the responsibilities of the public when they engage with medical human enhancement. Echoing an argument in the literature on the duty to pursue scientific research,²⁴⁶ there appears to be a duty to pursue the ongoing development and improvement of scientific understanding, which is hard to pin down, and which runs contra to bioconservative or techno-sceptic Luddite positions. Transhumanists correspondingly tend to suppose that human beings are naturally driven to pursue self-improvement strategies, but the basis for such a universal human duty to engage with enhancement and medical self-evolution is hard to identify without referring to generalisations about the value of human life and the content of wellbeing.

Hughes has suggested that one contradiction of transhumanism is that 'although most transhumanists are liberal democrats, their belief in human perfectibility and governance by reason can validate technocratic authoritarianism'.²⁴⁷ Echoing the literature on the duty to carry out scientific research, which prioritises autonomy and beneficence as key principles

driving best practice, a nudge approach to human enhancement as a public health project aims to maintain autonomy and variety in the lives of individuals, avoiding homogeneity without compromising the herd benefit. This approach aims to resolve any contradictions transhumanist goals may have retained from their Enlightenment roots, which still colour the public perception of the transhumanist paradigm.

Group projects sometimes involve action that requires a commitment from but does not directly benefit individuals or seem to reduce their own suffering. Their action matters for group success and ‘free rider’ behaviour would be undermining.

Using authority to govern on matters where individual autonomy is necessary, for example in the delivery of medical expertise, can be controversial. Dworkin argues that the use of authority is granted to public health endeavours because of the beneficial consequences of public health goals which would not be available to individuals acting alone. Public health, although not limited to the distribution of generally agreed upon public goods, is a protected category of interventions which does not include warfare, for example, to protect against the abuse and overuse of justifications of authority made by appeal to public health.

It does not seem that autonomy of individuals is automatically overridden when they act for reasons relevant to group goals, however. Some have suggested that a wider conception of autonomy is useful in the increasingly personalised medical paradigm because genetic information about individuals has power beyond immediate patient-clinician interactions. Dove et al,²⁴⁸ and Prainsack’s²⁴⁹ account based on solidarity conclude that a shift towards personalised medicine raises ethical challenges that can be better addressed given a notion of autonomy that allows the interests, wills, and support of others to play an active role, where the individual insufficiently describes how freely taken decisions come to be made. Personalised medicine means healthcare better suited to individual patients, aiming to ‘improve tailoring and timing of preventative and therapeutic measures’,²⁵⁰ using biological information including genetic information about patients and biomarkers.²⁵¹ Although the focus of personalised medicine is the individual, and has reinvigorated questions about the extent to which a more extensive consultation with the individual should be expected in clinician interactions, the detailed and specific information involved in increasingly personal healthcare also introduces a problem for group autonomy. The availability of precise and patient-specific genetic information creates ethical challenges, for instance where patient

autonomy conflicts with the implications of genetic information for others outside the patient-clinician interaction (for example in a separate patient-clinician interaction).²⁵²

Although it is unclear whether the use of genetic indicators constitutes a medical paradigm shift, because of the slow progress and complexity of diagnoses based on early biomarkers and genetic information,²⁵³ it seems, nonetheless, that increasingly many patient-clinician interactions have at their periphery the potential for discussion of heritable disease burden, based on better understanding of genetic information.²⁵⁴ Heritable genetic information about individuals by its nature can very often mean there is reason for more than one person to act, or for one person to act in ways that are likely to affect others.²⁵⁵ Because of the complexity of genetic analysis and the technological component of gene sampling, sequencing, and analysis, the information is not currently available or accessible to individual patients, who are reliant on the medical system as a result. Critics of increasingly technological and genome-focused medical care have warned that this reliance could give way to a pervasive technological paternalism. Individualisation of medicine drawn from impersonal analysis of genomic information held by clinicians could draw focus away from patient values and preferences and towards unfeeling clinical interactions. Whether a paradigm shift or not, this shift could begin to risk patient autonomy, without correspondingly advanced regulation and safeguarding in place.²⁵⁶

One likely avenue for medical enhancement amounting to transhumanism is heritable genome editing to reduce the risk of disease. Transhumanism has been criticised as an ideology that makes grand and unrealistic claims about the possibility of eliminating genetic weaknesses or predispositions to disease which are intended to be attractive, but are unlikely to be realised.²⁵⁷ However, medical shifts towards individual engagement with clinical decisions about risk, given predictable causes of suffering and preventable death in individuals and family lines is a much less far-fetched possibility for the future of medicine, and one which has been taken seriously by ethical bodies and medical institutions separate from transhumanist activism. The Chair of the 2018 Nuffield Report on Genome Editing and Human Reproduction identified as a major challenge for research into the ethics and best practice of HGE the unattractiveness of human editing:

“Because the moral acceptability of intervening in the human germ line with the aim of affecting the traits of future children is highly contested, often engaging deeply held views, I was acutely aware that, whatever conclusions the working party arrived at, some people were bound to oppose and even abhor them.”²⁵⁸

As CRISPR technology has been in the public eye since 2012 and, since 2018, is now known to have been carried out on human beings with some success,²⁵⁹ action is needed to prevent the accumulation of negative attitudes to HGE while the use of technology like CRISPR on human beings is still comparatively untested and more clinically risky. Human enhancement through genetic editing could create lasting improvements in lifespan and health-span, but would require widespread participation from consenting adult patients, and their consent on behalf of future generations. An account of autonomy that involves group responsibility and consideration of the consequences for more than the individual seems important, if such decisions are to be made well by laypersons in clinical settings.

One of the recommendations Nuffield 2018 made to the UK government was that ‘broad and inclusive societal debate’ should come before any decisive action was taken to legislate about the use of HGE.²⁶⁰

For non-experts to give meaningful consent to participate in complex emerging health issues on the global scale in conditions like these, when the long-term consequences of heritable editing are not yet clear, good public engagement is important as soon as possible. As also identified by the WHO in their official communications about HGE following the announcement about events in Shenzhen in 2018, engagement with the public in transparent and comprehensive global communication about ongoing work on gene editing has not yet happened outside academic circles. The future of human genetic editing, if left shrouded in mystery, seems more likely to lead to unease and resistance than a more transparent approach, but it is unrealistic to expect people in general to understand and appreciate what gene editing will do, or why it should be appropriate in the light of technological medical innovations that the health of the group should become increasingly important to the individual.

Similarities between the effects of the emerging technologies of transhumanism and other interventions on the body to augment and enhance biological capacities, for example birth control, demonstrate that the integration of enhancement with medicine is inevitably socially unpredictable and potentially risky, but that the risks can be offset by clearer deliberate communication about norms and aims of intervention. Provided individuals understand that the long-term consequences of research and development in general are not yet clear, but that the goals and norms involved in the application of developing technology aim at justice, some of the problems for enhancement medicine are reduced significantly.

A slightly limited definition of transhumanism as human medical enhancement described in the primer chapter allows discussion of duties to plan for the next likely transhumanist events in relation to species' wellbeing and global disease burden. Existing general medical practices are more consistent with human enhancement to the extent that they fulfil a limited set of aims for the enhancement of the two human core capacities, cognition and lifespan. Other technologies should not be considered transhumanist simply because they make life easier or healthier for human beings, which distinguishes the ethical problems of transhumanism as a public health issue relating to human enhancement, from laws governing the development of AI or surveillance technologies, or projects focused on inhabitation of other planets.²⁶¹

As described in chapter 3, current challenges for the ethical development of HGE include uncertainty over precision of outcomes using CRISPR-Cas9, and establishing who, if anyone, has the authority to change the genes of future generations of human beings.²⁶²

The Nuffield report gives a defence of HGE, judging the social value of enhancement uses of HGE against two measures:

- i) Nuffield sets aside 'innovation' as an unhelpful way to think of HGE, and suggests it is better to consider what kind of society would come from using an innovative technology like CRISPR, and what kind of society would want to use it.²⁶³
- ii) Nuffield measures the value of kinds of society in terms of reduction or exacerbation of social division compared to the status quo.²⁶⁴

Using these measures, the Nuffield report encourages certain applications of HGE, and supports research and application of heritable reproductive technologies including gene editing. Nuffield supports HGE applications only up to a point, however, because the argument based on these measures does not respond well to problems like the slippery slope, where the outcomes change gradually as the result of measuring against the current standard rather than what would have been acceptable to those assessing from a position of overview.

Alternatives to nudge, like education and legislation are costly, time consuming, insensitive to personal preferences, and need to be updated often.

Halpern identifies ideal behavioural insights policies as ‘low-cost and unobtrusive ways of nudging behaviour’²⁶⁵ and explains that the early work of nudge theory in public policy (in the UK at least) was to make surprisingly large improvements to the normal behaviour of people on uninspiring and mundane matters like loft insulation, paying tax, paying off fines, and insuring their cars.²⁶⁶ The benefit of the nudge, carefully orchestrated, is that there is minimal bureaucracy and minimal cost, compared to policies based on education or legislation.

In chapter 3, I introduced a problem for transhumanism, that even if bioenhancement to improve human lifespans, health-spans and cognition appears possible in clinical environments, this does not mean success. Much research now suggests that the technologies needed to allow targeted enhancements in individual patients are realistic and will eventually be ready for patient use. What is not yet clear, given the absence of consensus on applications of enhancement or about its consequences, is who is responsible for its management as a public-facing issue, beyond the global oversight of organisations like the WHO. Challenges relating to implementation of transhumanist projects would need to include a justification of the cost, as attempts to guide medical progress to include enhancement would very likely involve re-distributing resources from existing public health work into the delivery of therapeutic care.

Bostrom²⁶⁷ identifies the redistribution of resources as a problem to be overcome by technology through programmes of research and development. The urgency of the transhumanist goal to reduce the burden of senescence, to reduce the burden of death related to old age particularly, comes because the longer there is no way to address the disease burden of old age, more of that burden is passed down to future generations.

Through his dragon analogy, Bostrom tells a story about how negligible senescence can fail to inspire sufficient support to attract resources, because it does not aim primarily at disease as traditionally understood, so fails to register as an urgent burden to be relieved, rather than passed to future generations. Bostrom remonstrates with those who dismiss negligible senescence as a goal, and his conclusion shares much with Hughes’ criticism of pessimism about enhancement as Luddism, which would underestimate the potential of technological advances to change human life. Whether rationalised by bioconservative views or by concerns about technological paternalism more common to the political left,²⁶⁸ pessimism

about the potential of new technology to improve the condition of human life that is passed on to the next generation does not benefit anyone, say Bostrom and Hughes. In recent years, available funding for enhancement projects including gene editing and cell repair has increased significantly,²⁶⁹ but financial support is not the only barrier.²⁷⁰ The citizens of the kingdom in Bostrom's Fable ignored the possibility of removing the persistent problem, and the delay resulted in a smaller benefit for the next generation in terms of reduction of disease burden. Bostrom, and many other transhumanists, argue that the maintenance of the status quo is damaging and needs to be addressed at the level it does most harm: in people's attitudes to enhancement.

It is important to frame discussion of the difficulties of any duty to pursue enhancement, as well as attending to public attitudes to transhumanist solutions, in terms that are fair and helpful to the fears and immediate needs of the current generation, without unreasonably prioritising current over future suffering. Nudges, and behavioural insights more broadly, tend to be cheaper and less demanding to initiate than education, they put less pressure on existing resources because their primary function is to reframe and engage with public dialogue.²⁷¹

Choice architecture is inescapable: there must be a choice architecture guiding public attitudes to engagement with enhancement medicine and it would be better not to leave it to chance.

Legislation seems both inadequate and overly intrusive in personal matters, for example whether an individual maintains a healthy body weight for their height. Information and advice about managing diet and exercise exist in the public domain, including online resources and communities and government initiatives and, in many schools, physical education and home economics classes aim to instil relevant values early in life. However, levels of obesity in adults and children have risen worldwide in a general trend which began in the 1980s. In the case of body mass index, there are disparities across demographics which do not align neatly with, but often correspond in general with, minority ethnic groups and low-income groups, suggesting the likely cause of the increase is not a misunderstanding, or an unwillingness to manage lifestyle choices that contribute to body mass index, rather social and economic influences on behaviour.

Food packaging and grocery store layout (including online stores) cannot be avoided by most people and the physical appearance of groceries cannot help but create more and less attractive options for the grocery shopper.²⁷² The choice architecture of a supermarket or online store can be set up to prioritise sales of goods with high profit margins, to promote healthy choices, to keep customers shopping for longer, and so on. Choice architecture exists in public health decisions which happen privately, and the more salient and familiar options, those chosen by others around us, and options which present themselves as the status quo are more likely to be chosen, in general. The mere presence of expertise during decisions affecting individual and public health, for example advice from medical practitioners or access to information about lifestyle improvement, does not mean people always follow advice, form good intentions to follow advice, or succeed in maintaining their own wellbeing through consistently healthy and prudent choices. This is particularly true of choices which are not particularly salient, not the easiest or most popular, or do not appear consistent with one's previous behaviour at the time the choice is made.

Mongin and Cozic²⁷³ describe nudges as interventions that interact with rationality and rational failure of individuals at the time they make a choice, dividing the practice between three kinds of intended intervention acting alongside the existing choice architecture:

1. Nudges can redirect individual choices by only slightly altering conditions.

For example, a traffic light system on packets of sweets in a supermarket is more likely to result in individuals choosing the healthier option if the traffic light system has four, rather than three, coloured lights.²⁷⁴

2. Nudges can use failures of rationality instrumentally.

For example, nudges that aim to prevent littering or theft by using images of eyes or watchful faces rely on an irrational tendency to feel more under surveillance in the presence of eyes, no matter how obviously inanimate.²⁷⁵

3. Nudges can alleviate the ill-effects of failures of rationality

For example, a nudge can draw attention to an unconscious bias or a failure to reflect on the reasons to choose certain behaviour, as in the organ donation campaign which asked "If you needed an organ transplant, would you have one?"²⁷⁶

In a public health scenario where free-riders could pose a problem for group wellbeing, the second and third types of nudges seem less likely to be effective than the first type. Free

riders who do not participate for the group benefit deliberately, because they realise they stand to gain more as free-riders, seem unlikely to be swayed emotionally by a call to reflect on the selfishness or unfairness of their actions. Likewise, the existing bias present in the free-rider problem tends towards the self-interested choice, so nudges relying on this irrationality would not serve group wellbeing. Shifting the default, however, making a subtle change in environment can create a situation where free-riders must expend effort to avoid participation, no matter how small the effort is, rather than expending effort to participate.

The choice architecture for individuals encompasses biases, social expectations, physical architecture of the environment, and the metaphysical shape of the environment in which expectations are laid out in front of individuals. Not all nudges are alike, but the most successful and appropriate nudges in public health seem likely to be nudges that create a lasting choice architecture, either by promoting reflection on behaviour, or creating an environment where the best option comes more easily.²⁷⁷

I have so far raised three points in favour of an approach based on behavioural insights theory, to address ethical challenges facing the introduction of enhancement technology into general medicine. Particularly in large global jurisdiction where democratic process and individual freedom shape what is acceptable in public health policy and the use of authority in medical decision-making, the approach seems likely to improve the potential for enhancement medicine to promote flourishing.

In the next section, using an example from existing public health in the UK, I describe the motivations, implications, and outcomes of a project based on nudges, to improve medical outcomes and reduce the burden of disease through measures to engage positive public opinion.

Case from existing public health. Organ Donation in the United Kingdom.

The Organ Donation Taskforce was formed in 2006, with the main aim of increasing the number of available donor organs for transplant by 50 per cent.²⁷⁸ Organ donation in the UK was not compulsory or automatically initiated after every brain death and has not become either compulsory or automatic in the years following the campaign.

Donors after brain death (DBD) provide nearly all heart and lung, the majority of liver, and many kidney transplants. Maintaining a steady rate of DBD for kidney and liver transplants where live donation is also possible is important because the risk of death to otherwise

healthy Living Donors (LD) was given as 1 in 3,000 for kidney donation and 1 in 100 for liver donation in 2008. Between 2000 and 2006 the number of DBD fell from 739 to 633 and a startling increase in the need for risky LD was reported for both liver and kidney transplants (98 per cent and 284 per cent increase, respectively).²⁷⁹

The UK was reportedly one of the worst nations in Europe for organ donor numbers at the time the work of the Taskforce began. The Taskforce, by 2008, had identified the following challenges for improving the rate of successful organ donation across the UK: coordination and organisation (of organ donor services, including retrieval); legal and ethical issues; training for clinicians and administrative staff; public awareness and promotion of donation issues; identifying the role of the NHS.

In chapter 4, I explained why successful transhumanism must convince individuals that enhancement is something they should concern themselves with and should aim to normalise a greater degree of engagement with public health responsibility. Three of the challenges identified by the Taskforce for organ donation are particularly salient to the challenges for the introduction of enhancement medicine into the national, public, healthcare paradigm. The Taskforce planned to overcome the legal and ethical challenges of making structural changes to increase the reach of one national medical service; their proposed and enacted methods for raising public awareness and promoting recognition of medical issues included several nudge techniques involving messenger, group commitments, and defaults; and the important role of the NHS in the success of the changes they proposed for organ donation in the UK.

By 2008, 25 per cent of people the UK were registered as DBD on the NHS Organ Donor Register, which was an increase on previous years. Although public approval for organ donation had been reportedly as high as 90 per cent in the same period, with the vast majority of the population reportedly in favour of organ donation in principle, 40 per cent of relatives (in potential DBD situations) refused consent for donation of the viable organs of the recently deceased.²⁸⁰

The Taskforce observed, 'Organ donation is a 'local' activity, but transplantation can only be undertaken successfully as a UK-wide integrated service... Only a UK-wide service can identify and allocate suitable organs to meet the needs of these patients.'²⁸¹ One recommendation of the Taskforce in 2008 for the continuation of the project was that the NHS Blood and Transplant (NHSBT) service should become a fully integrated UK-wide service, to increase efficiency and raise the number of successful DBD.

Several ethical challenges were predicted for the changes proposed to the organ donation service, including possible conflict of interests between the clinical staff treating patients and the nation-wide NHSBT. The Taskforce also predicted and planned for differences in policy and procedure between local NHS organisations and National services like the NHSBT (as local organisation is more normal within the UK NHS system, and is generally preferred, to avoid clashes of policy).

Though 60 per cent of families consulted²⁸² were willing to donate the organs of a deceased relative, the NHS suggests a figure of 80 per cent of willing families would be needed to meet demand. If the figures for support of donation could be matched by the figures for actual uptake, significant improvements to patient wellbeing and emotional stress on recently bereaved families, as well as streamlining the duties of already busy consultants and practitioners who must acquire consent in each case, could be achieved. The Taskforce identified some risk of negative impact on families and patients resulting from such a 'radical change of practice' to make donation usual rather than unusual during end-of-life discussions and care.

The nudge which changed the default to opt-out registration was intended to tackle inertia and did result in more donations and a larger national resource for patients on the waiting list. The shift from understanding organs as a body part owned by the individual patient whose autonomy and interest were of primary importance, with donation as a rare event, to a national resource subject to political emotions like pride, responsibility, guilt, and commitment, is interesting because the shift occurred during a comparatively short time.

The change to opt-out does not prevent individuals from avoiding donation if they wish, and it does not impose financial penalties on individual who do so. The success of the nudge relies on a tendency towards the default, but also on presentation of information to increase the salience of the problem, to make people more aware of the need to remain on the organ donation register, and to reframe donor organs as a national responsibility and a source of national pride. Opinion polling is seldom completely unbiased, and the way questions are presented, the order, their phrasing, the questions flanking them in the poll, the approach made by the poll taker, the sample population approached, the perceived messenger, unspoken confounding effects during interview, for example, are well known causes of polling inaccuracy.²⁸³ When a poll fails accurately to represent the opinions of the public, it could be because 'the public' tout court were not consulted at all, rather a think tank or a select group close to the polling organisation. When the results of polls are organised and

analysed, unspoken confounding factors like squeamishness about death²⁸⁴ when confronted with a polling question (that would not affect reflective deliberation about actively pursuing organ donation but might affect quick decisions) might change the relevance of the number who indicated they were 'likely to donate' against those 'very likely to donate'. Another possible confounding factor in preference polling is the chance of individuals wishing not to seem extreme in their views, and therefore opting for a perceived middle ground that they do not occupy in private deliberations or in discussion with a group of like-minded individuals.²⁸⁵ Although the result of public opinion polls might not give completely accurate insight into the nuanced or reflective views of individuals, they do *take the temperature* of a more general public perception of a norm.²⁸⁶ What the individual on the street, presented with a personal medical question from a stranger, with little context, is willing to divulge about their support of a change in policy, might say a lot more and more accurately about the *sense* the public have about what is a normal view on the subject than anything else.²⁸⁷

Because the legislative change from opt-in to opt-out relies in part on the justification of favourable opinion polls, as well as on a corresponding programme of carefully presented information in favour of donation around the time of polling, the case involves the use of several nudges on the same target population, orchestrated to build and develop the attitudes that contributed to the success of the project, over time.²⁸⁸

There is some concern in the nudge literature that nudges should be narrowed to include only those interventions that do not attempt persuasion, particularly in interactions with more vulnerable social groups, because persuasive discussion engages reasoning. For this reason, some of the techniques that were employed in the organ donation case are less clearly nudges than others. A campaign that reminds individuals to register as donors through a prompt, for example with a simple message like 'Have you joined the organ donor register?'²⁸⁹ are more clearly nudges of salience than, for example, a poster advising 'Drop a drink size: 3 pints of beer is like having 2 burgers',²⁹⁰ which seems more educational, giving calorie figures and an easy comparison on which individuals could make better (free) dietary decisions.²⁹¹ Interventions that do not maintain the balance of liberty and paternalism that make the desirable behaviour easy, attractive, social, and timely without applying force seem more like propaganda, particularly if they aim to inspire negative emotions about other behaviours. To the extent that national pride in a shared resource might be a kind of irrational, though beneficial, belief that sustains social cohesion, aspects of the organ donation project could be scrutinised as a kind of well-meaning propaganda.²⁹²

Persuasive media messages based on shared norms, defaults, and commitments, and persuasive messages involving narratives of national identity are hard to separate in every case. The former seems a reasonable way to make sure policy is well-aligned with national feeling, while the latter might be uncomfortably close to propaganda, particularly if it motivates individuals to feel part of an 'in-group'. To justify the opt-in to opt-out change of legislation, the explanation was that people's nudged behaviour would better match their stated preferences if the default was changed. Justifying attempts to influence the strength of stated preferences around the time of polling relies on scientific findings about the object of the campaign. The implicit good of a society having enough donated organs available to prevent death and suffering appears to justify the use of emotive posters highlighting the risk of a shortage of organs.

Is the difference between public health and clinical conversation too great?

One complication for this approach to public encouragement without compromise to autonomy is the challenge that patient autonomy is more relevant to clinical interactions than to public health participation, and that the two are separate influences on individuals. A person could be reasonably and autonomously engaged by a public health project, armed with a message about social norms and justice, and yet still be imposed upon by technological paternalism in a clinical setting, when faced with a decision relating to enhancement interventions.

In predicting the likely impact of increasing the frequency and normality of organ donation, Taskforce²⁹³ proposed clinical safeguarding within healthcare trusts to reduce the likelihood that conflicts of interest between medical institutions would not adversely affect end of life care. One significant example of such a risk, given by the Taskforce, was that the pressure to secure viable donor organs might lead to end-of-life care being cut short prematurely for patients who appeared to be close to brain death, to expedite the release of organs for transplant.

Behaviour is not the same as attitudes and beliefs, and the Behavioural Insights Team report²⁹⁴ highlighted the 'intention-behaviour gap' between the stated preferences and self-report of previous behaviours given by individuals and their actual and likely future behaviours. The BIT Report identifies that self-reporting is error and bias prone and is not necessarily a good guide for people's feelings about their choices at the time the choice was

made. Relying on self-reporting or polling to identify the best way to guide behaviour, or relying on past behaviour as a guide to people's current attitudes, is not a straightforward matter of 'helping people do what they want to do', therefore.

In the organ donation case in particular, it seems that public health should not be aiming to perpetuate the social constraints that already maintain unfair and unjust differences in treatment and access to treatment. Social pressure to contribute already exists but this does not mean it should be made more a part of medical care or change the way clinical care is delivered. If anything, arguably the opposite should be the goal of good public health projects, to make people feel more empowered and less subject to social expectations placed on them when they make healthcare decisions for themselves against the grain of social norms and defaults.

When people are polled, they are not necessarily likely to report what they feel privately: in public and when responses are being recorded, people may say what they believe to be acceptable in public, to their peers, or to correspond best with the way they wish to be perceived, in combination with their own thoughts and preferences about what good behaviour should be. Relying on stated preferences and attitudes collected in this way seems more likely to perpetuate the status quo and to allow bioconservatism to persist in matters of public health than to welcome innovation and change. In the organ donation case, the justification for intervening on behaviour came because most people believed support for donation was the acceptable opinion to have, though they did not privately find it urgent or particularly salient. This dynamic would not hold in cases where the attitude that was perceived to be the most normal or acceptable in public was that novel gene editing technology is morally uncomfortable or risky.

The (financial) costs of Transhumanism as a project like organ donation campaigns: Risk given insufficient evidence for nudges

Although I will not spend very much time developing or responding to criticism of the cost-effectiveness of nudges, some critics of the claims of nudges and behavioural insights theory suggest that, although the light policy touch may be cheaper and less demanding of individuals or governments to implement, they are also less scientifically well-evidenced. Lower cost comes with less guarantee of effectiveness than legislation or education, and

nudges, because of their engagement with fine social differences and emotions, are more likely to go awry in unpredictable ways.²⁹⁵

In the organ donation case, the success of nudges in improving the survival rate on the transplant list in the UK is hard to judge against cost-effectiveness, in part because the Behavioural Insights Unit was formed only a decade ago and has undergone restructuring since. The cost of establishing taskforces, projects, and specialist teams to monitor changes in process within the healthcare sector also creates a threshold cost, and the financial benefits to societies of social change take time to materialise. Further, earlier reports from the Taskforce did identify the logistical challenge of coordinating a nation-wide organ retrieval service, with an accompanying ethical group on organ donation in clinical settings (clinical triggers for notification of the retrieval service, etc.) and the overlap of responsibilities and possible complications of competing policy or service provider schedules could yet reduce efficiency overall.²⁹⁶

Should irrationality be protected from nudges in cases like organ donation?

“Relying on heuristics, mental shortcuts, and intuitive decision making, real-life agents often systematically diverge from the traditional, abstract ideal of rational choice. Such divergence is a necessity and often works remarkably well. However, as proponents of nudging argue, real-life decision making is also rife with cognitive biases resulting in decisions that leave people worse off by their own lights.”²⁹⁷

Even if interventions to aid people with their decisions does not amount to authoritarianism, when authorities insist that people are irrational in pursuit of their goals and would benefit from being nudged, their intervention seems paternalist.²⁹⁸ Mongin and Cozic²⁹⁹ divide nudges that help people overcome irrationality to achieve better outcomes from interventions that help people use their irrationality to arrive at better outcomes. Although the former sounds more likely to engage individual reason and prompt self-reflection than a nudge which relies on irrationality to achieve the outcome desired by authority, this is not necessarily the case. If an individual is nudged to ‘overcome’ irrationality and arrive at a better outcome it does not follow that the individual is aware that this is what has happened, or that they would be able to reflect in order to repeat a similar pattern in a different context where the nudge was absent.

Although the goal of nudges may be to create a swell of behaviour that promotes reflection, social mirroring, and familiarity with a new kind of behaviour, it does not seem as though a nudged individual who has overcome irrationality in one case has done more than avoid irrational behaviour on one occasion. The possible harm of the nudge could be to bypass the exercise of 'decision-making- muscles' which could carry the individual at a later time.

Some nudge interventions allow people to remain irrational and make no attempt to improve decision-making processes, relying on rational errors like the automatic preference for the default to do so. The organ donation case is one example of a nudge that allows people who have a bias towards the status quo to arrive at a more socially responsible choice than they would have made, had the default been different. Interventions based on encouraging and solidifying the choice of a default option seems more likely to promote socially acceptable behaviour than to promote rational behaviour, if the effects of the nudge are long-lasting, therefore. If reason is not promoted by nudges that make use of irrational tendencies in this way, then the risk of reducing divergent behaviour would not seem likely to aid the uptake of new treatment options or innovative solutions to problems that exist in the status quo.

Social organisation and public support drive practice in clinical situations

So far, the success of the organ donation Taskforce seems to show that autonomy of families can be respected on the death of their loved one, where the patient was also an individual with autonomous decision-making power, and where the state has an interest in the decision they make. A wider group goal, for example the registry of donor organs as a national resource, does not undermine patient autonomy though it may guide the structures that support individual choice, and create the need for more extensive safeguarding in clinical decision-making.³⁰⁰

In response to an ethical concern, that increasing the frequency and changing the default of organ donation after brain death might create a conflict of interest between the healthcare system and individual patient care, the organ donation Taskforce recommended a study into the likely impact of a system of clinical triggers. The recommendation of the Taskforce was that clinical trigger points should be put in place to manage when the donor transplant coordinators were notified about a diagnosis of brain stem death, to standardise the process and create a normal approach which would be easier to monitor for irregularities. One

recommendation of the Taskforce was that the notification of the donor transplant coordinators should be automatic at a certain point in the process of diagnosis and treatment, regardless of whether the clinical staff attending to the patient believed donation would be viable after brain death. The intent was to create a common practice of assuming that organ donation after death was an option, rather than varying the care of individual patients because of the potential that their organs might be viable for donation.³⁰¹ This recommendation of the Taskforce as a response to the risk of structural injustice, when individual autonomy and group autonomy are in conflict, helps demonstrate how public health and clinical expectations interact positively to address concerns about technological paternalism.

Additionally, the recommendations of the Nuffield report on Genome Editing and Human Reproduction, because of its focus on the kind of society that would result from introduction of applications of HGE, is better placed to offer support to transhumanism than are arguments based on utility, or the expected medical benefit of individual enhancements.

Nuffield was an improvement on earlier defences of HGE in routine medical care in part because of a better account of wellbeing. Previous reports on the ethical use of HGE have suggested that the application of innovative medical technologies should, for example, 'promote wellbeing', which is both too narrow and too vague. Too narrow, because of its exclusion of otherwise good projects which failed to promote wellbeing from their inception, though likely to contribute to wellbeing less directly through a contribution to better practices, and too vague because the promotion of wellbeing is not explicitly described as either explicit promotion by campaign or promotion by successful visible improvements in wellbeing for existing people. It is plausible that wellbeing would not result after all in such a case, but that there might have been every intention and expectation of promoting wellbeing in intervening. The difficulty caused by the kind of social division that transhumanist technological interventions are predicted to exacerbate is that the technologies aim at an increased general wellbeing for individuals, and even for society in some cases, but do not support fundamental human dignity. This concern for fundamental human dignity, which may be grounded in human rights, includes resistance against homogenisation by collective action in one direction, or a slippery slope of pressure on patients in one direction.

A defence of the cost of nudge interventions, as the organ donation case is too recent to see long-term results

Particularly where projects are new or still in development, where they seem likely to change applicability or change quickly with new technology, the nudge approach seems preferable to reliance on the roll-out of education programmes or an abundance of fine-grained legislation. Infrastructure change and education programmes can be costly, so they should be accompanied by a change in public feeling which makes them more likely to survive and succeed. Although the financial cost of allowing public attitudes to novel technology to change in their own time is low, for example in response to market forces or overwhelming pressure from other nations, leaving choice architecture free to develop on its own seems risky. In chapter 4 I suggested that the introduction of personal computers is an example of the consequences for public engagement with technology that is introduced more quickly than expected, or without care taken to reassure and contextualise technology. Even more so than the introduction of personal computers in the workplace, the public health components of medical interventions seem more likely to cause anxiety and techno-stress on their introduction, if they remain familiar only to niche groups of interested specialists and private laboratories during development.

In the case of organ donation, most people in the UK were already happy publicly to express approval for organ donation procedures when polled. As a result, one challenge for the Taskforce was how best to motivate behaviour to match the general approval for transplants, to increase the number of organs available to the service via donor registration. This shift in behaviour, even given general approval for the medical intervention, took years to move from report to real changes in organ donor numbers, but the infrastructure that was created in tandem with public engagement, and on the strength of a public debate about national responsibility and effective donation as a reason for national pride, seems more sustainable.

Gyngell et al, responding to findings of the Nuffield report, support the moral imperative to pursue further development of heritable gene editing and applications of HGE to human reproduction. In terms of cost and the re-distribution of resources, they suggest, HGE is not only acceptable but imperative. The findings of the Nuffield report about HGE in particular were that it promotes wellbeing by reducing the rates of chronic, heritable disease and premature death in the population; promotes equality by lowering the risk of some people to match the risk of others; promotes justice by lowering health-costs for countries attempting to distribute resources.³⁰² Correspondingly, although it seems desirable to keep

the financial cost of the social debate low, accompanying research and development measures necessary to produce viable applications in healthcare, it seems justifiable in comparison to cases involving similar expectation of disease burden reduction.

Against preserving irrationality: choice-sets and prompting better reflection are more important

One claim often made by nudge theorists is that the influence of choice architecture is inescapable, so pursuing deliberate positive nudges is more socially responsible than allowing accidental nudges to guide behaviour. Nudges to improve the choice architecture often rely on identifying the reflective preferences of people, which can be undermined by automatic mental processes which favour choosing the default but without changing the choice-set.³⁰³ Making a change to the choice architecture to help automatic processes better suit reflective preferences does seem more ethical than allowing the choice architecture to promote undesirable outcomes by inaction, if the same choices are available before and after the nudge. In public health cases, the relevant reflective (considered, deliberate, rational) preferences at stake are usually good health and long life, which automatic (unthinking, accidental, non-rational, even lazy or biased) behaviour may not best serve.

The criticism that nudges only allow the choice architecture to allow the atrophy of decision-making muscles, promoting and re-iterating choice according to default, does not hold if nudges help people to reflect on their behaviour and the reasons they chose as they did. In public health cases, the good health and long life of the individual who is nudged might not appear to be immediately at stake, so persuasion is involved, to make better outcomes for individual health more immediate and attractive at the time of choice. In situations where commitments and habits are involved, for example in self-preservation or self-evolutionary activities, the urgency or salience of behaviour that is more likely to lead to the desired outcome might not always be obvious at the time that matters most.³⁰⁴

People need not be irrational to be ill-equipped for processing new kinds of risk which might not always present themselves clearly as urgent or salient at the time of choice. The future applications and off-target uses of enhancement are as yet uncertain, in many respects, and it does not seem irrational to be wary of unfamiliar medical interventions, nor does it seem necessary to suggest that people would be irrational to follow a nudge, if it seemed to support their values and preferences at the point of choice.

Where people are ill-equipped to judge risks adequately enough to escape the overwhelming choice architecture given reflective evaluation at the time of choice, support through management of the choice architecture seems more justifiable than in cases where a moment's reflection about the right course would reveal the right thing to do.

Nudge campaigns like the organ donation campaign match exactly the kind of considerations necessary for early social debate about the introduction of enhancement because they identify what kind of society comes from the positive choices of individuals, rather than concentrating on the differences between a possible society which has enhancement and the status quo. Nudges that shift the default to put obstacles in the way of less beneficial outcomes encourage individuals to act as if the personal risk were more immediate, by making the less beneficial outcome less appealing and less 'close' in the choice landscape. In the nudge literature, moving choices out of the default position is one way of making them difficult, unattractive, antisocial, or untimely. Anti-nudges of this kind are sometimes described as 'sludge' which lies in the way, between individuals and some choices.³⁰⁵ The response to critical concern that enhancement does not aim at a just or democratic society is that a just and democratic society does not leave sludge in the way of individuals contributing to a valuable and valued national resource.

In the case of organ donation, intervention encourages people to participate by making decisions that are unlikely to mean action in the near future, and which do not usually mean any action at all on the part of the participant. The contribution to the collective disease burden of participation is to reduce the overall risk of death of individuals on donor waiting lists by collective action. It would not be sufficiently helpful for 400 people to register as organ donors if the waiting list held 400 patients, for example, because of the gap between registration and donation. Changing the default makes the distance between registration and donation (between decision and outcome) benefit any emotional reaction the donor might have to the idea of their organs being transplanted, where it previously was a barrier to public health responsibility that made the issue seem less urgent.

The change made by shifting the default in this case does not impose burdensome costs (finance or time) on societies, particularly on the vulnerable, and does not mislead individuals by presenting any option as less risky or more urgent. Removing the sludge³⁰⁶ of the distance between 'now, when I have to do something to help' and 'later, when the problem might affect me personally', the organ donation project shows how people can be involved in the goals of a project peripherally, without increasing any burden on individuals,

misleading individuals about their contribution, or withholding information about the consequences of their participation.

Government action, including regulations about technology to manage risk can be justified both in terms of protecting existing people and the expectation that protection will continue to be of benefit to future generations. In HGE, the gap between action taken now and the benefits that might be seen by future people is greater than in the organ donation case. However, some applications of HGE would mean reducing a high risk for some people to the low levels already expected by others, which is easier to understand in terms of the organ donation case and similar existing nudges.

Respiratory disease is a known result of pollution from road traffic. There are strong moral reasons for governments to insist on green initiatives, and for the scientific community to develop exhaust filters, which can reduce incidence of respiratory disease by 40%. Air pollution is also a known epigenetic modifier, causing not only incidence of disease in the directly affected population but heritable mutations in their offspring³⁰⁷. If the target of the application of technology is straightforwardly therapeutic, addressing an imminent problem, for example by curing an existing disease or reducing its risk directly, the justification is not as strong as in the case where action now could reduce disease burden at a much later point. Heritable effects do not seem to pose an ethical challenge in this case.³⁰⁸ If the heritable effects were expected to be negative for subsequent generations, then there might be a conflict. Ethical objections based on our distance from future generations, or the burden of exhaust filters and green initiatives to existing people for future gain seems misplaced in part because of the continuity of other goods and resources when generations change.

Gyngell et al. argue that the moral imperative to act, given applications like the above, can be most easily drawn when there are therapeutic benefits to an existing generation. Given the therapeutic benefit to an existing generation of patients, any benefit to subsequent generations is a side-effect which can also strengthen the justification for intervention. Nudges can change behaviour in the short-term and need not be permanent, but they can also change social norms, which are longer lasting and can affect the norms and preferences of future societies. In examples given by Thaler and Sunstein, norms can be uncovered by nudges and nudges can begin to change behaviour overtly in a way that begins a snowball effect to change norms.³⁰⁹ Medical procedures to elongate and enhance life could become more normal, from pre-natal genetic testing and editing to fine-tuning later in life, and might not pose ethical concerns to future societies,

Similarly, as many have argued, ethical transhumanism should not be only for the more privileged individuals in wealthier nations.³¹⁰ The mere existence of enhancement technology does not guarantee its accessibility to those in need, and without accessibility could do more to foster division and disunity than to enhance the species. Group commitments based on emotional reasons to act should reduce the inequality of accessibility. In some cases, inequality of accessibility has been found to result from personal views of those in the medical profession, despite best practice guidelines that require professionals to be neutral regarding moral value judgements about approved medical procedures. Other social and non-medical factors have been found to affect distribution, with socio-economic reasons meaning some patients may feel uneasy about seeking out enhancement medicine.

Another way of understanding the concern about the influence of personal views on the uses of medical interventions, separate from the actual uses of medical interventions, is that the mere existence or legalisation of HGE technology in the medical sphere could make it less comfortable to exist as an individual with genetic imperfections, as judged by the prevailing norms.

Academic concern about the lasting effect of heritable gene editing on the autonomy of future generations captures the sense of a more generalised public concern that increasingly technological medicine is a threat to the freedom and wellbeing of individual patients and of the unenhanced who might begin to feel marginalised. Encouraging individuals to engage with public health projects so that public health projects better represent their views and needs is also important and could make the continuing global shift toward transhumanist medical interventions more socially responsible, less elitist, and less risky. Successful nudges would aim to reduce the extent to which individual choices about enhancement would be governed by concern about behaviour aligning with the views of others, even though people sometimes prefer to do things for poor reasons or on the basis of trends, and probably always will.

Conclusion.

The use of authority in some situations, where group action can achieve something unavailable to individuals choosing for themselves, is justifiable even when public health projects make interventions which restrict individual freedoms. The likelihood that

transhumanism, as medical human enhancement, will be such a project seems very high because of the focus on species self-evolution, heritable enhancements in future generations, and changes to the status quo of the human core capacities; intelligence and lifespan. In transhumanism, where large groups of individuals must act together, oversight by an authority by appeal to a consensus about what the public good should be, seems appropriate. It has been argued by critics of transhumanist human medical enhancement that there can be no such consensus about the good of the goals of transhumanism, and that such consensus is not found in transhumanism.

Nudges can be used in public health settings to offset generalised issues that are not necessarily urgent for the individuals at the time they are nudged, but whose contribution to a social dialogue are important for ethical progress. I have suggested that nudges can also be used to influence which public health contributions appear in widely accepted narratives about progress and individual responsibility, and so change which risks are considered tacitly to be acceptable by the public and which benefits are sources of national pride.

This chapter built on the argument of the previous chapter, suggesting that authority and autonomy are both involved when public health authorities aim to guide human self-evolution, and are justifiable in the face of technological uncertainty, particularly given the persistence of established narratives about human bodies. In this chapter I have argued that the approach taken by a behavioural insights-led series of nudges, which has been productive and beneficial in the United Kingdom in recent years, could be adapted to remove the stigma from human enhancement as medicine. Human enhancement medicine, in the form of the capacity to genetically modify human embryos for implantation and pregnancy, is now approaching human trials under increasing global scrutiny, in a wider global health context than had been predicted by earlier literature on the risks of Transhumanism. Transhumanism would be more ethical with the capacity to avoid more of the challenges it faces in homogeneity, elitism, and paternalism, if its implementation remained in a more egalitarian global health spotlight. The resources to manage transhumanism as a public health endeavour, rather than allowing private research to continue alone or secretively, seems worthwhile and must be accompanied by public engagement.

In patient-doctor meetings, in clinical settings, expertise is important to help patients reach good decisions and there are safeguards to ensure that patient autonomy is not weakened by the authority of experts. The individual patient must choose individually, even though they are guided by a clinician who knows more about the content of their choices and about

the risks and benefits at stake, and even though their decision is also likely guided by a family or social support network with their own preferences. In public health, where there is not expert consensus about the best option or where the choosers are not the experts, there are principles and safeguards to prevent organised collective action from becoming coercive or forceful. Consensus conferences, phased trials and peer review aim to prevent the undue influence of individuals or groups from biasing the direction of movement. Behavioural insights theory accepts that there is a choice architecture that exists in the space between public and personal healthcare consensus, where narratives about health and wellbeing that are ever-changing are driven by myriad factors beyond the reach of legislation but need not be left 'unattended' to drift into undesirable places.

In this chapter, I have described a recent case involving the legislation guiding organ donation legislation in the United Kingdom, how the approach of the public health campaign has changed over time, with public response to the early campaign efforts as one of its guiding principles. I suggested that recent action to improve public health by increasing the number of donors after brain death was not at risk of becoming a programme of forced organ collection because of safeguards maintained by public health authorities, even though the national organ donor network acted independently as a service provider, outside local authorities and separate from central government. The national organ donation campaign to encourage donation behaviour and raise awareness avoided becoming emotionally manipulative and was not coercive propaganda. The campaign used some recognisable nudge techniques from behavioural insights theory which are nevertheless controversial, in some ways capitalising on irrational status quo bias towards defaults, but increasing the salience of decisions for choosing agents, the outcomes of which would be urgent and present to other, distant, people.

The campaign asked individuals to make a decision that would likely benefit other persons after their own death (as the most common scenario for organ donation is donation after brain death). On this basis, because the campaign appealed to a shared commitment to the flourishing of the group, the organ donation campaign seemed sufficiently ethically similar to a possible future case of human gene editing in which individuals could take an interest and contribute by commitment to pursue species' goals rather than to fear medical interventions (*i.e.*, in contrast with personal hygiene, vaccinations, and other public health responsibilities with more immediate personal rewards).

The sudden emergence of HGE and human enhancement medicine into the public eye as the plausible future for general medicine has the potential to surprise the public and cause alarm, compounding existing tendencies towards apprehension about technological medicine. Transhumanism approaches fundamental conceptions of health and wellbeing through the lens of core human capacities which can nonetheless be improved upon, and it is insufficient to rely on assurances that enhancement interventions will be part of the healthcare service by the time they are patient ready. Technology that aims to change human core capacities seems existentially challenging in a way that could prime the public to react negatively to emerging enhancement medicine, particularly in the area of genetic editing. The ethically appropriate approach to enhancement medicine should not rush human trials and it would be naïve to expect universally positive responses to events like Shenzhen, given its ethical shortcomings. An ethical and democratic approach also should not allow inertia to contribute to apathy about enhancement medicine. Without deliberate and careful management of attitudes to enhancement, the high likelihood is that events like the Shenzhen trial risk causing long-term damage to the progress and perception of enhancement medicine.

The campaign did not ignore the role of emotions in collective responsibility for deaths on the organ waiting lists. The organ donation campaign also used emotive content, with messages targeting fear, guilt, pride, and commitment to shared responsibilities. This approach is not uncommon in public health, and in the next chapter I will examine the weight of emotional reasons in clinical decision-making about engagement with medical human enhancement as part of a complex social and politically charged public health project.

Justifying Emotional Nudges in Public Health

Nudges sometimes engage strong emotions to reframe issues as salient and urgent to individuals. Can this be truly constructive, given the need for democratic public debate about emerging technology and novel applications of medical technology?

Introduction

In A3 I asked whether, although emotions are not in conflict with reasons but are part of reason, the emotions that guide decisions about emerging technology might amount to less valuable reasons, for example if anxiety or disgust about the unfamiliar are categorically more likely to lead to worse outcomes via conservatism and stagnation. In the face of a need for people to undertake reflective deliberation about what they feel and what they prefer, it seems counterintuitive for the work to be done by nudges.

If it can be negligent or irresponsible to allow oneself to be guided by intuitions without reflection, opening the door for decisions made from bias without any attempt at rational consideration, it also seems acceptable that such decisions could ethically be modified by libertarian interventions. In B2 I argued that such interventions, rather than atrophying the capacity to reflect and develop reasons, can promote reflection, and protect against accidental harm from poor choices while they do so. Studies have shown that many effects of unconscious bias, for example social prejudices, can be overcome by prompting. Once the bias is identified, reflective self-inspection of motives and desires follows and, in many cases, the effect of the bias on choice vanishes.³¹¹ If deliberate framing of the issue could mean that the need to reflect differently were made apparent to the individual, it seems possible that behaviour would better match reflective preferences rather than unconscious influences, for example where the injustice of prejudice is apprehended, rather than merely understood.

Critics of transhumanism have asked what will happen to human emotional states when human bodies are enhanced, particularly if cognitive capacities are changed. I will focus on the following criticisms of human enhancement in terms of its negative emotional effect on decision-making, justice, and social cohesion:

First, the criticism that greatly elongated life, the absence of death, or negligible senescence of cells leading to greater robustness against ageing, would also lead to boredom amounting

to meaningless life. I will agree with conclusions from the literature on boredom and immortality, that only true eternity risks seriously undermining the meaningful activities of human being, whereas greatly elongated or otherwise improved life does not.³¹² However, I suggest the kind of concern that the boredom criticism of negligible senescence represents is also present in general attitudes to enhancement and contributes to an intuitive misapprehension of bodily suffering as valuable and productive. This may often be the case but does not mean suffering is important to a valuable or productive experience of life, nor that it should be favoured over the reduction of suffering from enhancement.

Second, the concern that enhancement would create division, undermining the shared and universal bond of 'what it is like to be a human being', that this shared bond is fundamental to systems of cooperation and so enhancement could undermine the foundations of justice and fairness. Hughes explains that libertarian transhumanists must engage with democracy, since 'only alliances with other cultural and biological minorities, and a strong liberal democratic society and state can ensure that posthumans are not persecuted'.³¹³ Critics of transhumanism have expressed concern about similarities to eugenics and argue that the unenhanced seem much more likely to be persecuted or to become undesirable outcomes of reproduction, in societies where enhanced individuals are the goal.

Criticisms of enhancement from arguments about emotional difference raise the problem that no matter the policy put in place to ensure democracy in delivery and distribution of resources, equity and security of minority groups, once the fundamental connection of human beings through shared core capacities is lost, justice is at risk. In such a situation, the pressure on individuals to become part of the 'in-group' would be too great for true autonomy in choosing enhancement.

Part of the reason for a lack of positive engagement with transhumanism is doubtless a general tendency to prefer the known over the unknown, and a preference to avoid unknown risks without social support, and for inertia. Though these tendencies can be a valuable and constructive part of reasoning, and emotional responses are essential for informed meaningful consent, they can also lead to irrational outcomes where emotion changes the salience and urgency of certain risks or goals. Sunstein explains variation between individual willingness to 'take the leap' away from social convention as something that is not tied to particular political views, but simply varies across populations, and in general the minority of people are willing to act unsupported by their peers, in most kinds of risk-taking scenarios, at least most of the time.³¹⁴ In A3, I suggested that the challenge for

policymakers would be to avoid putting too many people in a position where they must mask their strong beliefs to avoid condemnation from peers, but allowing that policy could still make an impact on social norms such that undesirable behaviour does not persist and ill-adapted views do not lead to stagnation.

When emotional reasons to do something out of the ordinary do not align well with an individual's reflective views, or where reflective preferences are not fully formed because emotional reasons for action appear more salient at the time, for example in matters that appear to be matters of straightforward preference like tastes, individuals should be responsible for engaging with their own reasons and preferences in an active way. In general, it is important that individuals are responsible for their own emotional states and preferences, rather than emotional reason relying on intervention from authority, but this does not mean that emotional reasons are off-limits to policy. The use of (libertarian) authority to improve the choice architecture is still justifiable in some cases that appear to be about tastes, for example about satisfactory conflict resolution, social security, and just treatment of others. In this chapter, I suggest that the use of strong emotion in public health campaigns is more acceptable when it draws attention to commitments and responsibilities to others and avoids categorisation as negative propaganda though it may be a kind of propaganda more generally construed.

Because of the likelihood that transhumanism as human enhancement medicine will provoke negative affective responses for many people, based on the existing understanding of technostress and technophobia, such interventions are ethically justifiable, in the way they have been justifiable for previous public health problems.

Background

How technology can be integrated poorly, an example from outside healthcare.

Biometrics have been used in schools in the United Kingdom for around two decades, with schoolchildren learning how to use their thumbprints for locker and library access, school lunches, and as a measure of attendance from an early age in many UK schools since the early 2000s. Legislation to govern the protection of databases to secure privacy of biometric data came later, in 2012.³¹⁵ Concerns over Human Rights violations³¹⁶ and the issue of parental consent³¹⁷ followed and, later still in 2007, the UK Government began to discuss stakeholder and external discussion of the application of biometric technology in schools, at

which point the practice was normalised in a large number of UK schools, rising again between 2007 and 2009.³¹⁸ The order in which the implementation of this application of biometrics took place was widely criticised as an attempt to lower the resistance that would be encountered by surveillance measures from a generation of youngsters.³¹⁹

As Agogo and Hess³²⁰ explain, the modern world 'is flooded with technology' giving people 'little choice whether or not to avoid it.' This can place stress on relationships with technology which can become focused on particular applications of technology and can quickly become controversies in cases like the use of biometrics in schools. When the order of the introduction of technology to society does not seem to respect the expectation that there will be democratic processes like consensus reporting and public debate, when the use of authority does not appear to value public opinion, or to require public engagement or awareness before proceeding with plans. The negative affect relating to unwanted interactions with technology can be severe for individuals who experience technostress and anxiety,³²¹ but the more mundane negative effects of unpleasant affective responses to technology can also be harmful and are experienced by many people as fear of technological change and protectiveness about the status quo.

Medical practitioners are not always keen for technological interventions or new areas of medical intervention on the body to be incorporated into the medical professional domain. The 'reluctant medicalization'³²² of abortion in 1960s Scottish abortion law reform was a reluctance against the effective monopoly on the practice of procedures to terminate pregnancy. Terminations had previously been a common law offence in Scotland but were effectively given over to the jurisdiction of medical expertise, wherein medical professionals could exercise the freedom of their best medical advice, for their individual patients, if the medical condition of the patient made termination the best course of action.³²³

"[I]t was possible for a medical practitioner, acting in good faith in the interests of the health or welfare of his patient, to terminate a pregnancy after a careful study of all the circumstances of the case, and after due consultation with appropriate medical specialists. Thus, abortion was only a crime in Scotland if criminal intent could be proved, a doctor having freedom to practise medicine in this type of case, as in all others, according to his clinical judgement"³²⁴

In an interesting footnote from Davis and Davidson's 2006 paper, the confusion and social tension about terminations for even reasons relating to the wellbeing of the patient are clear, with the continued anonymity and confidentiality of these statements going some way to

demonstrate how controversial and charged the termination of pregnancy was, in relatively recent decades leading to decriminalisation, and which to some extent remains present in abortion law in many jurisdictions globally.

“Oral testimony suggests that indications other than of an emergency medical nature were little used, most doctors believing social and psychological grounds to be illegal. Transcripts of a series of interviews with retired general practitioners, gynaecologists and psychiatrists, Apr. 2003 to Apr. 2004, are held by Gayle Davis. These were granted on condition of anonymity and confidentiality and are not available for consultation.”³²⁵

The social consequences and the harm that came from confusion about the scope of medical jurisdiction, because medical care was governed so severely by social consequences, affected patients and practitioners. Davis and Davidson describe one of the only criminal cases that was brought against trained medical practitioners for carrying out termination, during the 1960s in Scotland, explaining that the case was singular, and likely came about because the nature of the procedures was commercial and correspondingly lacking the standard of care that was expected of the medical profession, favouring financial profit ‘rather than being performed in good faith for therapeutic reasons.’³²⁶ The police were happy for the medical profession to act within their professional domain, and were not intent on prosecuting trained clinicians for carrying out their best judgement, nonetheless the stigma about termination led to medical reluctance, both to carry out the procedures and for the procedures to be designated to medicine.³²⁷

Open discussion about social taboos can be very challenging and cause great social division, as the abortion reform case shows. Early reform laws in the UK were driven by a few medical professionals who saw the enormous need for social change to reduce the suffering, and medical procedures which would allow ‘freedom from excessive fertility’³²⁸ and improve life expectancy and quality of life. Their proposed measures to carefully legislate to allow termination for social and psychological reasons were compared unfavourably with eugenics strategies, and concerns were expressed about ill-informed patients accessing services that would harm them, without fully appreciating what was at stake. Discussion about the reasons women would seek out termination of pregnancy was also heavily embedded with discussion of emotional reasons for choice, and emotional ways to dissuade patients from undergoing treatment, which persists in many jurisdictions today.³²⁹

A dialogue which is ongoing, and often involves persuasive and even coercive measures, is the issue of mandatory ultrasound procedures before terminations can be carried out. There

are now several US states where the US Supreme Court has upheld the law that women must see the image produced by an abdominal ultrasound scan prior to termination, and where clinicians must capture the ultrasound image and deliver a 'detailed description' of the fetus to the patient before the termination procedure. (See for example Guardian 2019, and other global reporting on the forced abortion ultrasound laws which are enforced in the US states of Kentucky Tennessee, Louisiana, Wisconsin, and Texas, as of July 2021.) Laws of this kind have been criticised for their cruelty and insignificance to medical outcomes but have been supported by various factions of the pro-life lobby in the United States. The more charitable reading of such laws is that they are an attempt to ensure the patient is fully emotionally aware of the consequences of their medical treatment, but this justification fails, partly because the procedure requires a trans-vaginal ultrasound, a forced additional procedure which is uncomfortable and invasive, and does not add information relevant to the medical outcomes of the procedure or the clinical expertise given by the doctor to their patient.

Consent in medical settings cannot be called informed in the sense that is important for patient wellbeing if it lacks the element of emotional appreciation for the consequences of choice, but the overinvolvement or the abuse of emotion in a clinical setting, as the 'heartbeat' laws demonstrate, do not serve to strengthen individual capacity for free, informed, rational consent. Charland identifies that an individual without the capacity to feel emotions would not be satisfactorily competent to consent, in the sense required of clinical decision-making.³³⁰ Such an individual, although they would be able to grasp that clinical expertise was relevant to their physical condition, might not be able to imbue meaning to choices, or value to outcomes.³³¹ In the case of 'heartbeat' ultrasounds, the legal requirement of procedures that are medically unnecessary,³³² misleading,³³³ and have been called 'psychological torture' attempt to force patients to feel certain emotions, to dissuade them from undergoing medical treatment. The legislation has led many clinicians to question the intervention of legislation into this area of medical practice, expressing similar views to those of Baird and the Metropolitan Police in 1960s Glasgow, as described by Davis and Davidson, that expertise in the professional domain should be allowed to proceed unfettered, within reason, once an issue has been designated to the medical domain. Deliberately engaging the emotions of a patient at the time of their decisions must be done with their medical best interest in mind, and should not be driven by social norms about, for example, the importance of the nuclear family or the teachings of any religious organisation. Nussbaum captures the spirit of the requirement that just societies should engage with emotions even though they drive a default, in her description of public art:

“Public artworks have to set a default option; the only alternative is to have no public art at all, or only art of such stunning mediocrity that it communicates nothing.”³³⁴

Although establishing a default in the engagement of public emotion is one of the criteria of good public engagement, because one of the stipulations of nudging as libertarian is that it should leave the choice set unaltered, it would be unreasonable for interventions addressing emotional reasons to rule out particular affective responses altogether, if they had not been unacceptable before. To return to the cafeteria example, the carefully organised display positions of the carrot sticks and fries could hardly still be called a nudge if the choice of fries meant diners could expect scrutiny about the emotional state that led to an unhealthy choice of fries, or intervention to insist that their choice of carrot sticks should be made cheerfully, unlikely as the eventuality seems in the cafeteria setting. When authorities nudge but do not prohibit, they should be careful not to make the affective state of individuals the focus of negative attention when they do not follow in the direction of the nudge, and not to pathologise or demonise dissent. An example can also be seen in the rhetoric of public debate about birth control and termination of pregnancies, where proponents of safe, legal medical provision of reproductive care identify their position as ‘pro choice’ rather than ‘pro abortion in x, y and z defined circumstances’.

Behavioural insights work in public health should not use emotions to make certain choices more vulnerable to peer criticism. In case policy interventions invite others to respond negatively, or where interventions risk changing the choice-set too much to be described as nudges, they no longer promote reflective deliberation about emotions and reasons for medical preferences, but rather increase the number of people in a position where they must mask their strong beliefs to avoid condemnation from peers.³³⁵ In such a case, individuals would be acting to avoid the negative social consequences of dissent, rather than because the choice architecture had enabled self-reflection about the connection between their immediate behaviour and wider values.

Nussbaum’s argument for targeting values and emotions rather than merely actions

Nussbaum places the capacity for emotions central to the requirements for maintaining a just society. Emotions in general, like fear, guilt, pride, and disgust, and the culturally shared values integral to political emotions more particularly, determine people’s commitments. Commitments made by individuals to shared goals increase the bonds of shared culture (*i.e.*,

those cultural bonds between individuals realised in poetry, music, public festivals, and theatre) and social cohesion. These not only create good feeling and individual feelings of belonging but also discourage people from treating others as means instead of ends when it would better serve self-interest. As the result of good feeling, a sense of belonging and participating in culture, and resistance against individuals becoming mere means, an emotional (loving) society becomes more egalitarian and more just³³⁶.

Nussbaum's account describes emotions as object focused and involving evaluations. Consistent with the contemporary cognitive accounts³³⁷ this definition does not set emotions apart from, but places them central to the work of believing, valuing, and deciding rationally. So shared cultural beliefs and values, with awareness of shared cultural beliefs and values, propagates justice because it encourages just behaviour between individuals without the need for authoritarian control or coercion. Justice, according to Nussbaum's account, is fair and equitable treatment of individuals, regardless of characteristics like race, gender, or social status, and regardless of their potential utility. Justice, as egalitarian treatment, and freedom from persecution, rely on individuals treating each other as ends in themselves rather than behaving differently to those individuals who do not further one's project, or who pursue different goals. Justice relies on individuals receiving fair and equitable treatment whilst being afforded the freedom to express dissent and difference of opinion.

The clash of libertarianism and egalitarianism, also a longstanding philosophical problem, is addressed by Nussbaum's account of justice, which also helps make sense of criticisms of enhancement medicine. In particular, the need for shared beliefs and values supports criticism on the grounds that distributive justice would not be best served by promoting research into enhancement, even if enhancement meant better healthcare and more individual freedom of body. Because of the potential for enhancements to change the core capacities of some individuals (their lifespans, health-spans, *and cognition*), the political concern about enhancement medicine is not that an ideal human nature would be lost or changed, rather that the enhanced and unenhanced might lose the shared emotions that allow for just society.

Given the possibility that the enhanced would experience different emotions from the unenhanced, it is also possible that the shared emotions of just society, the condition for just behaviour between individuals that does not need authoritarian control, would be lost. Given the need to promote justice in fair and equitable societies, there follows the concern

that enhanced individuals might no longer perceive their fellow citizens as capable of sharing artistic, cultural, and emotional sensibilities, and would act unjustly towards them as a result.

Bernard Williams wrote that the life of Elina Makropulos became pervasively empty and painful to her because it lacked meaning. This meaninglessness was caused by boredom as she struggled to live three hundred years as the same forty-year-old woman. Williams intended the example in support of his argument that immortal beings would all, necessarily, feel boredom and become tired of life in this way (called the Necessary Boredom Thesis). The truth of Williams' thesis would make death, if not good, then at least a comparable good.

This concern about enhancement contains but is not limited to the necessary boredom thesis, that a long life without change through ageing would be miserable and would lead to alienation from other human beings. The suggestion that boredom reduces or removes the potential for life to be meaningful has been addressed in an emerging literature on the philosophy of emotion, wherein boredom is divided between the experience of life as dull, temporarily, such that immediate joy is lost in the activities of the day, and the kind of existentially challenging loss of interest in life which seems more like the criticism of enhanced life. Fischer and Mitchell-Yellin³³⁸ have argued that many events in a very long, even truly immortal, life would not lose their meaning by repetition because of the innate value or comparative impermanence of other features of life and events, even when the life itself seemed permanent. There would always be a chance that some activities and experiences continued to be meaningful. It seems unlikely, for instance, that EM would become bored the experience of her own children growing into adults, or at least no more at risk than an unenhanced person. Nor does it seem likely that she would grow tired travelling to see historical artefacts which were decaying or were newly discovered if these things had been interesting to her unenhanced self, unless the conscious value of the things was tied to her own vulnerability. Bortolotti and Nagasawa³³⁹ have argued that situational boredom seemed to be the kind afflicting EM in her repeated exposure to similar life experiences. Though EM tired of things seeming the same, there is insufficient reason to suppose that a life untroubled by cancer or loss of mobility would contribute to habitual, pervasive boredom, and loss of meaning in the Makropulos opera. Realistically, the experience of the things themselves could remain meaningful as could a gradual shift in EM's overall desires, not on the basis of bodily change through ageing but through new experience of the external world. For instance, having become a great opera singer, fully realised the potential of that career, felt its success, it is hard to imagine EM would have no interest at all in mastering another form of musical performance in reality, but fictional accounts portray

deathlessness as existentially troubling, and the narrative persists in conflict with more specific preferences about the value of long life and palliative care to reduce pain from age-related decline. There is huge variety in the responsiveness of individuals to boredom, in particular to situational boredom, but also perhaps to tendency towards habitual boredom, and there have been links drawn between boredom proneness and various mental health issues such as addiction, poor judgement of risk, and depression, for instance.

Critics of bodily enhancement agree³⁴⁰ that even if the alienation of the enhanced from meaning is not certain, any chance of such a split between the enhanced and unenhanced is too great a risk to justify enhancement.³⁴¹ Sparrow suggests that the pursuit of human enhancement can only eventuate in a 'rat race' where generations and demographic groups compete and aim to outdo each other, causing unjust distribution of resources and goods. It appears that either consequence of such a loss of shared humanity, from divisions either of preferences an experience of meaning, or from the brute competition of the enhanced and unenhanced would reduce the quality of life of many and would not be desirable.

I have suggested that nudges are a promising solution to the difficulty arising from enhancement as people must often be persuaded to behave in socially responsible ways, guided by their rational preference for the ethical pursuit of public health in a democracy. A further complication for transhumanism is that enhancement might create an incentive for some to succeed and diminish the flourishing of others who prefer not to engage with enhancement. In the literature on technophobia and technostress responses, two affective states towards technology³⁴² predominate. The first is that technology can improve life, take away tedious or painful tasks from human beings and reduce suffering, giving more time for leisure. The second is that technology represents a risk to human superiority, could undermine the unique capacities of human reasoning and ingenuity, and take jobs away from human workers who are not trained to find employment elsewhere.

In chapter 5 I argued that emotional constraints constitute real barriers to individuals receiving good medical care, particularly for vulnerable groups or those with marginalised religious beliefs that influence healthcare preferences but might not be widely accepted as 'good enough reason' to diverge from medical advice.³⁴³ I have also argued that emotions should be the subject and target of public health nudges in cases where emotional reasons are involved in maintaining the status quo, just as much as actions should be the target of nudges where action is undertaken unconsciously or out of habit. In the case of organ donor behaviour, the public health issue of organ shortages was improved not only by changes in

infrastructure to make donations more likely to succeed but by a sheer improvement in the number of donated organs, once the public both knew and felt that organ donation was important to them personally. Numbers began to rise before the legislative change from opt-in to opt-out was enacted in 2020.³⁴⁴ In the case of enhancement, it is not only important that people understand and appreciate how enhancement could be personally beneficial (salient and attractive to them, personally) but also that apparently salient and attractive propositions relating to novel technology are not a source of anxiety.

A further concern, given the evidence that the dark side of affective responses to technology seem to be experienced by a significant proportion of populations, is that division between people who experience technostress about medical interventions and those who do not is inevitable, and that some individuals simply will not flourish as life becomes more technological. Affective division about enhancement could become particularly damaging if enhancement, like personal computing technology, reached the point where its use was ubiquitous, inescapable, and necessary for individuals to participate in activities of everyday life.

Divergence of political emotions that change accessibility, rather than the divergence of core capacities appears to pose a risk to social cohesion. The result of divergence of affective responses, in matters requiring general agreement for social order and cooperation, risks either increased authoritarianism to recreate public order, or else public disorder and non-cooperation from low interference and emotional divergence. So, aiming to avoid a serious divergence of political emotion in response to enhancement through the absence of a deliberately beneficial choice architecture also means work to avoid the damaging social problems that might be predicted as the result of disagreement about technology, more than division caused by living alongside those with different capacities.

Nussbaum's argument that shared political emotions in general, and love in particular, matter for justice draws a distinction between the justifiable and the less justifiable ways governments might aim to influence the feelings of individuals.³⁴⁵ The work of governments to promote certain kinds of shared values and shared emotions, such as condoning and promoting pride in the anti-racist views celebrated on Martin Luther King Jr. Day, for example, is justifiable. According to Nussbaum's distinction, this promotion of value and emotion does not attempt to control matters of personal preference that should be acceptable to others in the society. In contrast, government promotion of the doctrine of a

particular religion would aim to promote one acceptable value or preference over other socially acceptable values and preferences and would be unjustifiable.³⁴⁶

Emotional divergence between the enhanced and unenhanced seems likely either to increase division or necessitate an increase in authoritarian rule and forced homogeneity of behaviour. In previous chapters, I have aimed to demonstrate that the outcomes of public health campaigns can be significantly improved with emotional nudges that balance liberty and benevolent paternalism, while ignoring emotional choice architecture for decisions about new technology can leave a lasting discomfort with technology in general that hinders creative implementation of possible medical enhancements in particular. As a consequence of accepting Nussbaum's argument, that emotions and shared values are important for maintaining justice and that some emotions and shared values are justifiable targets of authoritarian intervention, it is still important to demonstrate that the emotions and values involved in resistance against opportunities for medical enhancement are similarly justifiable targets of intervention, and that the lack of cohesion in emotions about enhancement would be politically divisive, bearing the risks to justice that Nussbaum identifies.

In a case addressed by Nussbaum, the divisions to be healed are structural and prevalent. Nussbaum draws out the difference when a society that marks Martin Luther King Jr.'s birthday, annually to uphold a commitment to tackle racism, deliberately does not celebrate or hold up certain religious belief systems over others.³⁴⁷

Nussbaum³⁴⁸ has also suggested that intervention can be justified, to catalyse a kind of healing and cohesion through feeling where societies encounter alienation and division. She argues intervention is needed in the United States, where one solution could be the reintroduction of compulsory national service, to bring a generation of young people together.³⁴⁹ Compulsory service, given a shared common trial to overcome, would better all young people regardless of social class, race, gender, or political views. One further aim satisfied by national service in particular is that a generation of young people, across other demographics, might interact with a more diverse group of people under shared conditions of struggle. As a means to overcome endemic problems like racism and homophobia caused by ignorance from simply never experiencing diversity, creating shared feelings, love, and justice even in this way does seem attractive.

Although the suggestion of national service is well motivated, in aiming to foster shared values and awareness of shared values, intervention to create and enforce a new state mandated obligation has seemed less justifiable than other less authoritarian programmes

involving education and incentives to good behaviour. Insofar as it aims to smooth out inequalities and prejudices through mandatory participation in a group project, the suggestion of national service goes further than a nudge account of behaviour management could support, as the first action taken. However, it is not in aiming to address wider social responsibilities through emotions and values (commitments to justice, to love others, to treat others as comrades), that participation in national service seems overambitious. Rather, that the shared values involved in military service are not universally respected or accepted in the way responsibilities and preferences for health and wellbeing are respected and accepted on reflection.

Nussbaum's suggestion that shared feeling is necessary for just society seems correct, and interventions to foster shared feeling relevant to the issue at hand are important for public health as for social justice relating to structural racial inequality.³⁵⁰

Would nudges guide people away from their own beneficial intuitions about medical bodily enhancement?

In chapter 5, Heritable Gene Editing (HGE) was described both in terms of its transhumanist potential, as a means to enhance human core capacities, and as an emerging perceived threat to our species' wellbeing. As an existing medical case in need of global by regulating bodies, and national legislation to contain the work of independent laboratories, as well as a kind of success for the transhumanist project, HGE is philosophically interesting. Judgements about what can be ethical and acceptable in the development and use of enhancement are already being made, globally, about HGE practices, and these decisions will shape future enhancement medicine. Public reactions generated about HGE in the wake of an announcement about its first documented use on human beings coming to light in 2018, included concerns about taking control of our own species' evolutionary bus.³⁵¹ I have mentioned the difficulty that can come when technologies are applied without public engagement with a new application, or with a known application in a new setting. The use of biometrics in prisons was accepted as a control measure to offer social security and to manage prisoners but was not readily accepted in schools, where parents felt that they should have been consulted as primary caregivers and should not have been effectively surprised with the revelation that their children were familiar with biometric scanners as a way to access learning facilities.³⁵²

Because unlawful and unregulated human testing rightly causes panic, confusion, bad feeling, and mistrust of new applications of beneficial technology, many negative responses to HGE technology can be expected if human trials are not approached with care. Shenzhen demonstrates that erroneous events also risk prompting negative affective responses that misdirect public debate about enhancement. Because the risk of misuse of enhancement technology is not negligible, not all fear or disgust is unwarranted, and this might seem to make deliberately encouraging optimistic attitudes seem misguided and even distasteful. In justifying the application of the same nudge systems defended in B2 about the organ donation case to the global transhumanist case of HGE, this chapter has argued that global feelings of disgust and anxiety are not often defended by individuals on reflection or about particular applications of the technology, are not beneficial to public health, and can therefore be justifiably nudged.

McConnell and Kennett³⁵³ address the claim that disgust and repugnance about bodily interventions and political views indicate a kind of 'deep wisdom' about human experience and human meaning which cannot be described by reason but captures the fundamental sense of what is importantly human and worth protecting.³⁵⁴ The arbitrariness of protecting some natural phenomena and not others, and the apparent short-sightedness of assuming that all human beings conform to one view of what is natural and that such judgements have not changed over time, in terms of behaviour and social organisation, are two reasons to reject this claim about the role of disgust in morality. McConnell and Kennett present a third reason, because Kass proposes that feelings of repugnance based on deeper wisdom about human life are distinguishable from general unease about interactions with the unfamiliar.³⁵⁵ McConnell and Kennett reject the claim that the two feelings of discomfort are so easily recognisable and suggest that the evidence from psychological studies into social behaviour demonstrates how easily moral judgements can be influenced by environment and other incidental factors.³⁵⁶

Even supposing the existence of a deeper wisdom that disgust and repugnance make available to us when confronted with certain events or technologies, it is possible that such wisdom could nevertheless be masked or confounded by environmental factors like warm cups or untidy desks. The complexity of neuroscientific explanations of feeling, on this reading of the experience of emotion as a signal of moral truth would still allow that there was something to be protected and attended to at the heart of human experience. A negative affective response to, for example cloning or face transplants, might still signal that truth even though its effects were less likely to be felt by individuals who i.e., felt financially

stable, were holding a warm mug, or had been otherwise primed to feel more comfortable in the moment, by confounding factors.

However, this description of the role of emotions in evaluative decision-making does not follow from the increasingly well-supported understanding of affective responses to stimuli as the result of a combination of social cues, reference to physiological states of the body, and reason-tracking heuristics from past experiences (memories, but not necessarily consciously reflected upon).³⁵⁷ Though basic emotions that inform complex decisions about safety and morality may be rooted in deeper inherited characteristics of societies or hierarchies, this inheritance is also traceable to primate dominance hierarchies, and does not seem sufficient to make a judgement that an apparently deeper feeling is the better guide to moral truth despite social change.

When presenting policy decisions to the public, it has been suggested that the role of emotions in evaluation is diminished when a choice eliciting emotional evaluation is presented alongside a choice requiring a cognitive mode of evaluation.³⁵⁸ This kind of joint presentation of options seems salient to public health concerns, where more than one mode of deliberation about the best course of action operate at the time a decision about, for example vaccination, organ donation, genetic enhancement, takes place. Joint presentation might also diminish the role of emotions like disgust and fear in the evaluation of novel applications of technology, because of the engagement of higher cognitive emotions and reflective reasons alongside intuitions.

Evidence from neurological and behavioural studies does not support the view that there is a deeper and more significant feeling of aversion to morally harmful stimuli which, although it may be misdirected by a warm coffee cup, holds the fundamental truth of the matter, nor does the strength or force of intuition remain stable. In decisions where there is a collective as well as a personal outcome to be considered, joint presentation of issues act as a nudge on behaviour that does not restrict choice or deny emotion altogether but makes a strong negative affective response less likely.

Misapprehension of transhumanism in shock reporting and overemphasis on technology as an intervention on the natural.

Bostrom and Ord explore the consequences of a general bias against changes to the status quo. The experiment asked participants to judge the moral acceptability of actions taken by

an authority, where the cognitive capacity of a large population was at stake and found most were content for authorities to intervene positively, to maintain the status quo against negative effects, but were less happy for authorities to intervene positively to improve on the status quo.³⁵⁹

In the case of medical interventions to change the core capacity of human beings, the consequences of this generalised bias in favour of the status quo are clear. Treatments which change the capacity of individuals to maintain the status quo tend to be considered more acceptable than enhancements, which also change the capacity of individuals but beyond the range accepted as normal. This difference is addressed in particular depth in the philosophical literature on cognitive and moral enhancements, where the concern is that even an entirely positive (i.e., very low risk of off-target effects, unlikely to be abused by future societies or governments) intervention to improve capacity might undermine identity or agency. An example of this kind of preference against enhancement despite well-understood risks and benefits can be found in discussions about sporting and competitive events, where the concern is that some athletes would have an unfair advantage over their unenhanced colleagues.³⁶⁰

In many cases of human enhancement research targeting health-span and lifespan, the most immediate benefit for individuals is reduced risk from the diseases of old age. Clinical trials to reduce the incidence of heart disease, diabetes, kidney disease, and obesity in old mice, mean enhancement achieved by genetic editing seems more like treatment to preserve a bodily status quo than enhancement. Much of the anti-ageing research carried out, including gene editing with the most potential to lead to human enhancement treat illnesses rather than making radical bodily changes. Because ageing and age-related loss of function are not the result of a single failure, a single weak organ, or a problem with a single gene, recent attempts to reduce and limit the damage of ageing which results in disease target multiple genes at once. Approaching anti-ageing medicine as a 'combination therapy', targeting two or more genes known to be linked to deterioration in one treatment, has already seen some success. The Harvard Church lab. Have discovered that targeting two of the three genes identified had positive consequences, reducing heart disease, kidney disease, diabetes, and obesity in mice, though combining all three therapies at once had detrimental effects. The team concluded that although all combinations are not necessarily beneficial, "holistically addressing aging via gene therapy could be more effective than the piecemeal approach that currently exists."³⁶¹

This change in approach, holistically targeting the genetic causes of ageing as combined therapy, to increase the health-span of mice, is also likely to increase lifespan, and Church identifies the research as a milestone toward reducing the likely side effects that result from administering multiple separate gene therapies, or from conventional treatment for multiple age-related conditions via medications.

Fear and disgust about enhancement medicine tend to be about individual isolated enhancements carried out on animal test subjects or by amateurs and are bound to be more likely to cause the kinds of emotions that make enhancement feel generally unappealing.³⁶² Messages from the media about the surprise and excitement of technologies like self-enhancement 'kits' or apparent leaps towards deathless life (pig heads kept alive without bodies, and similar reports of experiments with body-shock news appeal) could be moderated with messages akin to the behavioural insights that reframed tax payments, or the feedback reporting that increased open support for women in the workplace in Saudi.

In the latter example, the results of private polling found that the majority of adult men in Saudi Arabia believed that women should be free to join the workforce but also believed that the majority of adult Saudi men felt differently, *i.e.*, that women should not work. When the true beliefs of the demographic were reported back to participants, the study found the number of women at work began to increase.³⁶³ In the former, example, the message 'you are in a small minority who have not yet paid their tax' was more effective to motivate payment than a message suggesting non-payment was a wide problem, with many more culprits than timely taxpayers. By acknowledging the norm, nudges can motivate and, in emotionally charged cases, given appropriate targeting, can also help people understand that their fears and preferences are supported and shared by their peers.

Conclusion

In this chapter, controversy over the use of technology in reproductive technologies was one example from public health that highlighted what is at stake when an individual's emotional apprehension of emerging medical practices and social norms exists in tension with their other values and preferences. The literature on technostress and anxiety about emerging technology offers an explanation of what can go wrong when technology makes technological the previously analogue, or reshapes social differences as controllable, surveyable, divergences from desirable behaviour, as in the use of biometric fingerprinting

to track school attendance of 11- to 15-year-old children. The professional domain of medicine can expand to include the treatment of non-pathological conditions across a large subsection of the population and shrink to exclude from pathology human characteristics. Changes in the scope of medicine need not mean the characteristics, preferences, and experiences of individuals become homogenised or vilified, but there is a risk attached to such changes in medical domain and categorisation of pathology which should not be ignored by governing bodies.

The introduction of birth control to the domain of general medical practice was a source of tension at the time, as many GPs resisted becoming 'dispensaries' for non-illness prescriptions, and many patients felt uneasy about seeking birth control from their GPs. I suggested that the introduction of birth control into the domain of general medicine, as well as access through designated Family Planning Clinics is comparable with the inclusion of HGE into general medicine. It seems likely that applications of HGE will remain part of a separate and distinct field of interventions into enhancement medicine while early clinical trials, but true accessibility requires generalisation. The introduction of birth control into the professional domain of general medicine reduced some issues of accessibility, though imperfectly, causing other tensions for patients and practitioners, particularly those in small and close-knit communities.³⁶⁴

Some tensions about the introduction of birth control were specifically concerns about undermining social norms and the social order, *i.e.*, conservative concern about preservation of the nuclear family, the relationship between GPs and small communities where rumour and personal quarrels are more likely to arise, and the intersection of medical freedom and religious beliefs about moral behaviour. There are significant and useful parallels between the challenges for reproductive medicine in the 1906s and for the reproductive medicine that is likely given the advent of human enhancement medicine through HGE and negligible senescence interventions. Birth control as a medical possibility precipitated a change in the relationships between women and society, and between patients and medical practitioners. The freedom from 'excessive fertility' was freedom from a burden of normal biological function, which carried significant emotional baggage. Reform came about despite tensions over the health consequences of early birth control for individuals, and fears for social cohesion from a break-down of the family unit, facilitated by the move of reproduction into the medical domain, where clinical expertise could better act as counterweight to social pressures. Resistance from practitioners and society reduced not only the benefits and

effectiveness of birth control, and the potential for patient wellbeing, but also the accessibility of healthcare to certain vulnerable patients.³⁶⁵

Much like emerging genetic enhancements, reproductive medicine as birth control and abortion can improve the wellbeing of patients who are not considered unwell, and treatment can become necessary rather than elective because of patients' deliberation about the content of their lives, and their psychological needs. Introducing enhancement as part of the medical domain seems likely to mean fairer access to services than if independent enhancements were carried out by designated clinics. Critics of transhumanism are right, that social values will shape engagement with technology and that there will be divisions of opinion, but this does not automatically mean loss of social cohesion or the capacity of a society to be just.

In public health, where group cooperation is required for positive outcomes, differences of accessibility and differences of clinical experience across social demographics show how interpersonal relationships and social norms govern patient autonomy, as well as individual preferences and feelings. Changes to the shape of the medical professional domain, for example to facilitate better delivery of controversial medical care, can meet with resistance from the public and professionals when technology rather than social norms appear to be driving the introduction of new medical practices. Persistent bias that stigmatises medical care, if left untreated, can easily reduce the enthusiastic uptake of beneficial treatments, and can lead to injustice particularly during periods of fast social and scientific change where myths and misconceptions are more likely to colour social norms.

Thesis Conclusion

I have aimed to demonstrate that the outcome of public health campaigns can be significantly improved with emotional nudges. The very recent case of organ donation shows how an idea can become widely accepted, over time, to the point where more invasive (legislative) nudges are accepted.

I have aimed to show that the absence of a strong emotional framework, to guide reflection, leads to lasting damage where individuals are not able to address their biases. The consequence of a strong emotional framework, accepted within society, for individuals to arrive at their own decisions about engagement with technology is the first important step towards a transhumanism where individuals are not alienated by changes to their own bodies and core capacities.

I have argued that the justification for intervention also helps avoid the potential public health crisis of enhancement fragmentation, and that these measures also reduce the need to worry about particularly transhumanist public health issues like the boredom of immortal and enhanced beings.

With this justification to intervene and make emotional nudges, transhumanism can be an ethical project.

If patients have reason to resist medical aid because of deeply entrenched attitudes against enhancement, the individual wellbeing of patients and public health will decline, and the decline is harder to address after attitudes have become well-established than before applications begin to become salient to laypersons as decision-makers. Where social pressures make a treatment less easy, attractive, social, or timely, it is unlikely people will seek it out, even if clinical trials show the treatment is safe and effective. One solution is to address negative beliefs about the novel applications of technology, and another is to change the choice architecture to make 'seeking out' beneficial options less difficult.

I have discussed the importance of making the right interventions at the right time, to avoid undermining autonomy in the pursuit of enhancement. In this chapter, I have argued that some commonly held negative beliefs about the value of medically enhanced life, which have been supported by philosophical accounts of meaningful life, could present a problem for ethically acting on the goals of transhumanism to nudge people to choose enhancement. Lingering public mistrust in public health authorities and public health measures, as

encountered in the vaccine and some organ donation cases, suggest that the already vulnerable tend to suffer more when negative perceptions about safe and effective new treatments persist.

The pursuit of transhumanism does not and should not rely on wishful thinking. Transhumanist projects need not ask people to will themselves into being enthusiastic about something they fear and dislike and thereby make it more acceptable.

“The talk of believing by our volition seems, then, from one point of view, simply silly. From another point of view, it is worse than silly, it is vile. When one turns to the magnificent edifice of the physical sciences, and sees how it was reared; what thousands of disinterested moral lives of men lie buried in its mere foundations; what patience and postponement, what choking down of preference, what submission to the icy laws of outer fact are wrought into its very stones and mortar; how absolutely impersonal it stands in its vast augustness,—then how besotted and contemptible seems every little sentimentalist who comes blowing his voluntary smoke-wreaths, and pretending to decide things from out of his private dream!”³⁶⁶

Transhumanism does not guarantee nor require that everyone will choose the same way. The project of general reductions in the suffering and disease burden of natural ageing and vulnerability can succeed if *sufficient* people engage, similar to organ donation cases. Particularly in terms of public engagement about legislative change and meaningful development of norms, sufficient engagement means likely improvements in interpersonal responsibility relating to personal health as a part of public health.

“Our reason is quite satisfied, in nine hundred and ninety-nine cases out of every thousand of us, if it can find a few arguments that will do to recite in case our credulity is criticised by someone else. Our faith is faith in someone else's faith, and in the greatest matters this is most the case”³⁶⁷

Continuing progress relies also on legislative change and some public awareness of generally accepted guidelines. The legislative change of opt-in to opt-out in UK organ donation policy required some justification from public opinion polling and was an attempt to shift norms of public behaviour reflexively.

“Evidently, then, our non-intellectual nature does influence our convictions. There are passionate tendencies and volitions which run before and others which come after belief, and

it is only the latter that are too late for the fair; and they are not too late when the previous passionate work has been already in their own direction.”³⁶⁸

This kind of guided self-transformation is entirely compatible with the aims of the social and political ambitions of the transhumanist movement: provided autonomous individuals understand the implications of their individual contribution to a broader project.³⁶⁹ Additionally, this seems to address other projected problems for greatly extended life, relating to boredom from lack of interest in the state of other people and stagnation. Overt advocacy of personal responsibility for decisions made and attitudes adopted is what maintains the libertarian side of the paternalist-libertarian nudge approach to social transformation.

Though none of the above, relating to independent choice of exciting or stimulating projects as a better way for transhumanism to engage emotional responses, necessitates that the transhumanism case is more like choosing an ice cream flavour than an ethical code. Although individuals are unlikely to make decisions about enhancement based on their private study or analysis of scientific evidence it is not a matter of hysteria or getting swept along in beliefs with no basis. The best approach, mirroring other public health projects that encourage and reward participation is one that can attune public health responsibility and the values that make responsibility possible more closely to the changes that are happening.

It is neither desirable nor necessary for success that everyone chooses the same, but it is essential that societies share roughly the same values, as we might find in evidence grading panels, consensus conference reports, and organ transplant committee decisions. There is the worry that forced unification, although it is instrumental in some areas it is burdensomely overtaxing, unreasonable or unethical but in many contexts, including panels where shared assumptions and values are an important starting point, continuing without an agreed framework of some kind is much worse than waiting to discover one. A framework is important to help people understand medically enhanced life, to act justly, and to avoid mental health pitfalls like guilt, boredom, and loneliness (alienation) that would otherwise undermine transhumanist medical successes.

In this thesis, I have compared the work involved in changing the normative status of ageing and age-related decline to death, given the medical possibility of human enhancement, with other examples of efforts to change norms of behaviour and good practice in public health, most notably to attitudes to organ donation. I have framed transhumanism as a paradigm shift in medicine, where the core capacities of human beings, of longevity and cognition, are

matters of public health and can be improved by cooperation of the public and experts, guided by activists. The transhumanist approach, treating ageing as a disease and human capacities as open to self-transformation, requires a significant shift in public awareness and different responsibility about ageing, genetics, and death, which cannot be left to develop organically. The thesis suggests that authoritarian implementation of transhumanism and transhumanist values is undesirable, but that a libertarian paternalist 'nudge' approach is an ethical way to increase the likelihood that the inevitable increase in human enhancement as medical good practice is an ethical endeavour.

Making public health nudges to change attitudes, to favour the uptake of transhumanism as its ideas become more prevalent in medicine and global approaches to membership of species, is more ethical than making no interventions at all. The application of nudges was introduced to help resolve the ethical difficulty, identified in the philosophical literature on human enhancement as a contradiction arising from transhumanism. Transhumanism has elements of both libertarianism and authoritarianism in its aims, which result in a contradiction about the introduction of transhumanist practices to non-transhumanist society. When genetic enhancements become part of general medicine, their use might seem to risk increasing the homogeneity of the species, or to risk forcing individuals to accept enhancement, with the potential to undermine bodily autonomy. I have discussed, in the chapter 6 discussion of computer anxiety, the importance of making the right interventions at the right time, to avoid undermining autonomy in a particular area of transhumanist work. Because it is the transhumanist goal that is the most longstanding, and also seems the most likely to come about, efforts to reduce human susceptibility to ageing and illness, elongating natural lifespans and health-spans, by means of genetic editing are the most philosophically exciting part of the broad transhumanist project.

In chapters 4, 5, and 6 I have drawn comparisons between transhumanist goals, emerging gene editing technology, and existing medical cases. In chapter 5 in particular, I showed how carefully managed public awareness campaigns can help individuals change their own behaviour for the public good, to make smoother general public acceptance of a law to further support desirable behaviour. A further goal of the transhumanist in the medical sphere is that general medicine, transformed by genetic enhancement, would eventually give way to a more personalized medicine, dealing with individuals instead of one-size-fits-all treatments, but the contradiction raised by Hughes seems to suggest that individual autonomy in treatment by governing bodies is under threat by technocratic changes to norms and preferences about the use of technology on bodies (and in everyday life).

Chapter 5 considered an apparently simple recent case, where a nudge was used to improve public health. Organ donation law in England became opt-out, rather than opt-in, with the aim of increasing donor numbers. The mandate for making nudge interventions which might infringe individual autonomy by changing the 'choice architecture' came from establishing that there was a public interest at stake, and that in general people acting against the public interest were not doing so by design but by inaction and inattention. When the best interest of the individual also benefited the population, the justification for a nudge to change choice architecture increased, and when the attention of the individual was focused on the problem, their preference for organ donation tended to become stronger.

Postscript About Covid-19 and the Time of Writing

In the 1968 film *'What's so bad about feeling good?'*³⁷⁰ a contagious airborne virus spread through the population of the city of New York, carried in the respiratory system of citizens who felt no ill-effects or consequences other than a feeling of permanent elation. The government, realising that general elation would result in social upheaval, diminishing revenue from alcohol and cigarettes, and glum acceptance of the status quo, hurriedly enact a system of controls and restrictions to prevent infection. Restrictions to slow the spread of the virus included social distancing, handwashing and mask wearing, and an attempt to catch the wild toucan which initially had been responsible for carrying the virus around the city. The New York City authorities took these measures to try to control the spread of the virus, to protect the economy, and maintain the status quo. For contemporary readers, elements of the plot may bear an eerie likeness to measures taken during the 2019 global outbreak and spread of the coronavirus 'Covid-19'.

'What's so bad about feeling good?' is a farcical comment on government interference in public health issues. As the drama unfolds, the focus shifts from the threat of illness, as the virus carries no unpleasant or painful symptoms, to the threat of being controlled. In its portrayal of shady officials acting to 'protect' society against the disruption of feeling good, the film sent up the motives of authoritarianism, behind even well-meaning public health measures, wherever those measures place the group ahead of the individual and take measures to control individuals to that end. In the fiction, strong measures were taken by the authorities only when a change in public health seemed to threaten the economy, and the happiness of citizens seemed secondary. Having set the scene, the film follows the efforts of a group of radical free thinkers who resist the blithe cheeriness resulting from infection, and who wish to maintain their autonomous wilful scepticism, bitterness, and doubt. Eventually, the group are infected, and the authorities use them as guinea pigs to test a vaccine.

In the recent Covid-19 crisis, the virus does not cause elation, causes suffering, organ damage, and death, and its long-lasting negative consequences for the health of those who have recovered is not yet understood. Despite this, the reach of government authority, to curtail freedoms and impose sanctions, has often been the subject of concern, derision, and revolt (the latter has been demonstrated in particular by citizens of the United States, whose libertarian individualist ideology is perhaps stronger than in the United Kingdom, and is

certainly more prevalent than in countries like China or Singapore. The vaccination rate has been correspondingly lower in the US, despite comparatively good availability and advice about vaccines). The lack of clarity about the spread of the virus and the measures being taken has also been the subject of scrutiny by experts, the media, and the public. That certain advice about mask-wearing and physical distancing were endorsed by the World Health Organisation was not enough to dull the injustice felt by many, whose concern was for their civil liberties, which seemed to be at risk. Certain scientific advice, about how best to protect individuals against illness and death, could not always be prioritised over the need to protect livelihoods and the economy. This perceived tension of values often did little to reassure the public that the new rules and interventions were in their best interest, and it seems likely that later analysis will show the tension and mistrust undermined the effectiveness of lockdown measures.³⁷¹

There was precedent for demanding large-scale public responsibility in the face of a public health crisis (mask-wearing, distancing, hand washing), from the 1918 influenza outbreak in Europe and the Americas. A century later, comparisons were drawn between Covid-19 measures and the measures taken (and the behaviour demanded of the public) during the earlier pandemic, as well as to the famous spirit of national pride and cooperation (the 'wartime spirit') shown during the Blitz. Comparison with earlier national crises may have helped to ease the public feeling about more draconian Covid-19 safety measures and punishments for breaking the rules, but there is as yet insufficient evidence that the comparison to wartime spirit was widely accepted by the 2020 populous of the UK, let alone whether it motivated an increase in social responsibility.

When the UK government acted in the Spring of 2020, with measures leading up to the full 'lockdown' in mid-March, the likely behaviour and psychology of the public were discussed, and seemed to be considered, by SAGE,³⁷² and by members of the government's advisory team in their media appearances. Behavioural insights of the kind introduced by the BIT to UK government around 2010 were involved in predictions about public reactions and were widely criticised in the media as a poor excuse to delay lockdown. During the first lockdown period and in the period of cautious re-opening of non-essential services over the summer of 2020, severe financial penalties were introduced for individuals breaking curfew and rules were introduced to punish and exclude individuals refusing to wear a mask in enclosed spaces like shops, on public transport, and in public buildings. Later, regional action was taken to close businesses, clubs, and schools for weeks and months at a time, based on the available data about the R-number (the rate of infectious spread between individuals).

Mention of the behavioural insights phenomenon that individuals might become 'fatigued' by the lockdown, and gradually lose their resolve to behave according to the new rules was re-addressed in public communications about public policy, when justifications were given for maintaining or relaxing measures at particular points. It seemed to have been accepted that, in order for public health to benefit from even the best laid policy, the public must agree that measures were worth taking, and agree that the government shared their values. It appeared important to certain decisions that the public agreed that the virus must be controlled, that they agreed with the kinds of measures required, and agreed that they should take individual responsibility in the matter to protect themselves, their health service, and others. When the first news of a vaccine was announced in early November 2020,³⁷³ experts were cautiously optimistic, but stressed the need to manage public expectations, and to consider factors beyond the sheer effectiveness and existence of the vaccine, like distribution, priority of the most vulnerable, resistance against vaccines from some, and the need to persuade the public that a vaccine 'protects others as well as oneself'.³⁷⁴

This thesis was completed during a period of upheaval that was likely to be only the beginning of global Covid-19 related disruption. The thesis recommended nudge theory as a means to change public attitudes and behaviour in relation to a pressing global public health issue, namely the advent of large-scale transhumanist human enhancement technology. A multi-disciplinary analysis of the different actions taken by governments to control behaviour and manage expectations during the Covid-19 global crisis has been underway in sociological and anthropological contexts since the spring of 2020 and will likely be the subject of ongoing psychological and political research for years. The global response in 2020 may serve as an interesting example in the philosophy and ethics of medicine and public responsibility, though it was not feasible to revise the thesis to incorporate and reflect new and ongoing early research into the impact of behavioural insights and nudges on the way Covid-19 was managed in the UK. However, events at the time of writing do seem to provide further circumstantial evidence that uncoordinated attempts to *create* new behaviour quickly, in the absence of some consensus on public cooperation, comprehension, or agreement, result in worse public health outcomes than coordinated interventions that aim to influence preferences as well as actions.

Notes

Notes to Thesis Introduction

¹Though some public health efforts may only require the participation of a minority of the total population to be successful or may require the contribution of certain minorities (for example privileged minorities may contribute more to food drives, or for example a group with a certain rare blood type may be more called upon to donate blood, even if individuals within the group never happen to make use of these services themselves).

²Thaler and Sunstein 2008

³MINDSPACE report from the Behavioural Insights Team 2010

⁴See Appendix Fig 1.2 – 1.7 for examples of campaign posters to encourage donor registration.

⁵Though I have not included any images in the Appendix, campaigns against drink driving are often particularly gruesome and explicit, and have tended to draw attention to death and visible injury.

⁶Reuters September 2021 reported that France had plans to suspend medical staff who were not fully vaccinated against strains of Covid-19. This is consistent with French law about vaccination in general, where public school attendance is also predicated on child vaccines.

⁷In both the NHS Blood and Transplant 2021-22 strategy for donation in the UK, and the ‘ten-year vision for organ donation and transplantation in the United Kingdom’ for the plan ending in 2030. NHS Blood and Transplant 2021, NHS Blood and Transplant 2021a

⁸Randhawa 2011

⁹NHS Blood and Transplant 2021 pp. 34

¹⁰In particular, the Obama-Biden administration, 2008-2016 in the States and the Conservative-Liberal Democrat coalition 2008- DATE in the UK were interested in behavioural insights (nudge) theory as outlined by Thaler and Sunstein. Cass Sunstein is based at Harvard and, in 2020 was appointed Chair of the WHO technical advisory group on Behavioural Insights and Science, and Richard Thaler was involved in the creation of the Behavioural Insights Team in the UK. Thaler and Sunstein’s *Nudge: Improving Decisions about Health, Wealth and Happiness* is the first full account of behavioural insights theory that this thesis investigates in relation to transhumanist issues. Although nudges had been introduced to UK policy making by David Halpern and others in an earlier instantiation of the Behavioural Insights Team, which later became known as the ‘nudge unit,’ behavioural insights did not become fully involved in UK policymaking until around 2010.

¹¹For example, by creating a separate and carefully organised taskforce, a separate ethical organisation to judge specialized cases relating to novel technologies and applications, and to manage clinician-patient interactions relating to emerging or controversial technology.

¹²The Nuffield Report, 2018, advises that separating judgements about technology from judgements about certain ethical applications is best, though Ball and Holland suggest that such an approach may lead to difficulties for evaluations of emerging technology. Ball and Holland 2009 pp. 15

¹³Nuffield 2018

¹⁴Wellcome Global Monitor 2019

¹⁵See also, UK Parliamentary research which has recently begun into genome editing and the ‘future of food’ May 2021 (Parliamentary Office of Science and Technology), investigating the possibility of policy reform to reclassify Genome Edited food so that they are not considered GMOs.

¹⁶Regalado 2019

¹⁷Interest in bioenhancement and negligible senescence startups is rising, with supporters of Aubrey de Grey’s project SENS (Strategies for Engineered Negligible Senescence) reportedly raising \$5 million for the organisation in 2017.

¹⁸For example, in Bavelier et al. 2019 pp. 33

¹⁹Particularly, but not limited to the work of Sartre, Camus, de Beauvoir, and Arendt.

²⁰Two CRISPR scientists responsible for the discovery of bacterial ‘genetic scissors’ for gene editing, Emmanuelle Charpentier and Jennifer Doudna, won the Nobel Prize for their work, in December 2020

²¹At least 36 other organisms, including mice. Smolenski 2015

²²Ledford 2016 pp. 17

²³And the WHO have issued new recommendations about human genome editing as at July 2021.

²⁴Bostrom 2019

²⁵See, for example, Cavaliere 2019 and Tuerlings 2019 reports into ethics and governance of human genome editing, commissioned by the WHO after the Shenzhen announcement.

²⁶For example, about the status of behaviour carried out by individuals who had been cognitively enhanced, for example as a treatment for an eating disorder, or about the bodily autonomy of individuals whose parents had determined their likely characteristics pre-natally with genetic screening. In the case of eating disorders for example, research into genetic (heritable) causes means enhancement to produce physiological treatments for psychiatric disorders is not unthinkable, and the potential to reduce the incidence of causal factors for behavioural problems seems to be increasing. Habermas in Courtois 2019 pp. 452; Himmerich et al 2019 pp. 13

²⁷Pilsch 2017 pp. 8

²⁸Bostrom 2005; Sandberg and Bostrom 2006

²⁹Though, we should be wary of a certain kind of blithely optimistic futurism that has made some longstanding transhumanist goals appear less plausible, for example by enthusiastically predicting miracle advances and panaceas to solve human problems, as 'just around the next corner', since at least the 1950s.

³⁰Fukuyama 2003

³¹Arendt describes this as the 'human condition' rather than human nature, which is a convenient departure from views of human life based on biology alone, for the purpose of transhumanist discussion, and does not overemphasise the importance of the natural. Arendt 1998

³²Mill 2001 (1859) pp. 55

³³As behind Rawls' veil of ignorance

³⁴Gyngell et al. 2019

³⁵The cognitively enhanced could, purely by virtue of enhancement, think and feel differently about what constitutes social cohesion or personal autonomy, or about what kind of life was worth living. By definition, the morally enhanced might also have a different grasp of the importance of personal responsibility and social cohesion, and these effects need not be involved in the justification given in this thesis, for more urgent action to ease the implementation of more straightforward bodily enhancements.

Although, this is not to say that more bodily enhancements avoid the problem completely. Laurie Paul's vampires (Paul 2014) demonstrate that when a choice is personal, there might not be the option of recourse to evidence, and it might not be possible to apply personal experience to evaluate all outcomes. In this respect, Paul's suggestion that such change can be transformative, and that transformation poses a problem for rational decision-making, applies to many medical scenarios where momentous decisions are required of individuals lacking relevant experience, who may rely on testimony. So, transhumanism can require that individuals make decisions about their own personal transformation, without full access to their preferences and reasons in an enhanced state. Setting aside moral enhancements, the changes involved in transhumanist bodily enhancement involves epistemological transformation, bringing new experience and new information, and personal transformation, changing the lived experience of being oneself. This is not to say that transhumanist enhancements undermine humanity or morality, or at least no more than any more than other small-scale personal decisions would. Where Paul asks whether it is possible for people to make rational decisions where such transformative experience is at stake, the transhumanist must also ask how to engage the public fairly, about deciding to live a radically different kind of life. For this answer fully to support ethical transhumanism in practice, only pressure from the outside can help or hinder the individual in their decision, and this pressure must also be examined.

³⁶Griffin 2001 pp. 310-311

³⁷And, on the political left, another kind of bio-Luddite who does not value and is 'suspicion about the products of the corporate consumerist machine'. Hughes 2002

³⁸For example, about Mill's suggestion that "Neither pains nor pleasures are homogenous": Nussbaum's discussion of Mill's ideas about the variety of valuable modes of life, which considers the position of Mill's ideas on a point of fundamental tension, about whether human happiness was equal to pleasure (between the work of Aristotle and Bentham). Nussbaum 2004 pp. 65

³⁹Sparrow 2019 & 2015, Danaher 2019

Notes to Chapter 2: Transhumanist Primer to Human Enhancement

⁴⁰This stipulation, that there should be clear success in mice *and the potential to move to human trials*, defines one part of the Methuselah Mouse Prize, or 'Mprize' for oldest mouse. The Mprize, which aims to encourage and reward research into negligible senescence, offers a considerable sum: the prize money was \$10m (over £6m) in 2003 at the time it was first awarded. The prize awards the researcher or team who can present the oldest living mouse and, separately, awards the researcher or team who can demonstrate rejuvenation of ageing mice. The first winner was awarded the prize in 2003, seven years after its inception for longevity. Dr Andrzej Bartke won the Mprize with a mouse (unsentimentally known as GHR-KO 11C) who lived 1,819 days, or almost five years. Bartke had previously identified genetic variations in dwarf mice that resulted in 50% increases in longevity (2001). In 2004, the rejuvenation prize first went to Dr Stephen Spindler, who extended the lifespan of mice by 15% using calorie restriction in middle-aged mice, also extending health span and reducing cancer incidence. DNA microarray analysis showed rejuvenation in the mice (they became younger as the result of treatment in their middle age). These were, in many respects, early attempts, and the most recent winners of the prize have achieved results with genetic enhancement. Similarly, Aubrey de Grey as instigator of the prize, has emphasised the likelihood that future prizewinners, and the research most likely to contribute to human negligible senescence, will result from genetic changes rather than lifestyle interventions like calorie restriction.

⁴¹In support of this assumption see, for example Kurzweil, Church and de Grey as proponents of different aspects of posthuman and transhuman predictions about the future. Kurzweil and de Grey have suggested that there is very little question about whether a substantially augmented human future is likely, and Church has suggested that transhumanism is already a reality for many human beings because of technologically augmented life and technological medicine.

⁴²More about this in chapter 6 when I discuss activism in healthcare reform and emotions.

⁴³Mandip and Clifford 2019 give a brief history of human gene editing research.

⁴⁴He Jiankui was associate professor at the Southern University of Science and Technology, Shenzhen (China). Along with two colleagues, He is now likely to face jail time as the result of the Shenzhen CRISPR trial in 2018.

⁴⁵Editing of somatic cells is already in use to target cancers, for example. Church in Skerrett 2015

⁴⁶Lander 2016

⁴⁷Regalado 2019

⁴⁸Greely 2019 explores the ethical failings of the consent process at the Shenzhen CRISPR trial, one of which was the inclusion of a clause, agreeing a fine for leaving the study. pp. 164

⁴⁹By 2021, the WHO position remains that the wider implications of germline editing are yet to be evaluated by global stakeholders, and no agreement has been reached about costs or benefits.

⁵⁰HIV in particular not a good target for CRISPR-Cas technology in its early stages because of the other possible therapies available which do not require germ cell editing.

⁵¹The WHO launched its registry on human genome editing in 2019.

⁵²de Grey 2007 pp. 7

⁵³de Grey 2007 pp. 1

⁵⁴Bostrom 2005a

⁵⁵Bostrom 2005a

⁵⁶Zealley and De Grey 2012

⁵⁷The Church Lab Harvard Website

⁵⁸See for example, Kass 1985

⁵⁹Davis and Davidson 2020; Aniteye et al. 2016

⁶⁰Though the normal function account is not widely defended in the literature on enhancement.

⁶¹See Juengst 1997

⁶²Which was removed from the DSM-4 in 2017. See Kamens 2011 for a broader history of gender and sexuality in the DSM.

⁶³See, for example, Parens (2013) who explains the concern that the medicalisation of sociological problems is inherently bad, though rejecting the simplicity of the view and arguing that distinguishing medicalisation from over-medicalisation is possible, but complicated.

⁶⁴And the danger of enhancement to our accomplishments is explained by Sandel thus, 'The more we become masters of our genetic endowments, the greater the burden we bear for the talents we

have and the way we perform. Today when a basketball player misses a rebound, his coach can blame him for being out of position. Tomorrow the coach may blame him for being too short', Sandel in Herisson-Kelly 2012 pp. 204

⁶⁵Church has suggested that, given appropriate further research into its effects, heritable editing of germ cells can be safer, more efficient, and with less risk of off-target effects that would likely result when editing millions of somatic cells to address the same problem. Skerrett 2015

⁶⁶For one example, relevant and adjacent to public health policy, consider mobile applications that allow personal activity tracking and contact-tracing, which are the default on many smart phones. These applications track movement (logging an estimate of daily exercise, sleep hours, distance and speed of travel from a 'home location'), and proximity to other users, with significant implications for public health, once users allow access to the data via an NHS mobile application. (For up-to-date user information, relating to the most recent large-scale use of automatic tracking data, see for example the NHS Apps Library - NHS COVID-19, through the NHS UK website.)

⁶⁷Coin and Dubljevic 2020

⁶⁸Church 2017 pp. 5770-5771

Notes to Chapter 3 Introduction

⁶⁹Mackay 2021 in The Herald Scotland Online interviewed Eagleman about the future of such enhancements for general use.

⁷⁰Wolf 1997

⁷¹See for example discussion of Brannmark in Metz 2013 pp. 45

⁷²Putnam 2012 (1981)

⁷³Williams 1973

⁷⁴As part of the second dialogue, I will also acknowledge the three prominent philosophical positions on the difference between enhancement and therapy. This explanation will be brief, offering only my justification for reference mainly to disease-based and sometimes to professional domain accounts of therapy, to pick out enhancement as philosophically separate from therapy.

Notes to Chapter 3: First Dialogue

⁷⁵Saghai 2013, pp.491

⁷⁶Hansen and Jespersen 2013

⁷⁷Thaler and Sunstein 2009

⁷⁸UK Government's Behavioural Insights Team: MINDSPACE Report 2010

⁷⁹Menard 2010 pp. 233

⁸⁰It is also valuable for governments to know when to de-incentivise damaging behaviour that has become inadvertently attractive: an intervention which does not have the same coercive implications.

⁸¹UK Government's Behavioural Insights Team: EAST 2014 – Easy, Affordable, Safe, Timely

⁸²Ramos 2009

⁸³For example, at the Utah Valley University Campus.

⁸⁴For example, on the Hamburg Metro, where stairs near an escalator were painted as a running track.

⁸⁵Huang and Baum 2012

⁸⁶The up-to-date MINDSPACE (2019) report from the now quasi-autonomous Behavioural Insights Team suggests that in some cases incentives are less effective than other motivators. Carefully presenting information about the rates of obesity and the dangers of a sedentary lifestyle as a nudge to encourage people to join exercise clubs might be more efficient and effective than would be financial incentives offered only to obese individuals to encourage exercise.

⁸⁷A later chapter will consider with more depth the justification of practices behind a recent change in UK organ donor registration law, which was driven by the popularity of behavioural insights policy-work.

⁸⁸Mill, 1977 XVIII pp. 223

⁸⁹Mill, 1977 XVIII pp. 260

⁹⁰Joseph Raz (1982, pp.89) distinguishes between the relationship between political power and liberty in three traditions of liberalism; those where liberty is protected by reason, with rational

justification needed for political action against liberty; those where there are basic human rights to some liberties, such as freedom of expression, which are absolute and formally recognised (Lockean tradition); and those for which principles that justify private action do not necessarily apply to political action, and so protect liberty of individuals (the tradition of J.S. Mill). The latter two traditions better recognise that political interference is less justifiable in some areas of private life than in others.

⁹¹Mill, 1977 XVIII pp. 223

⁹²Freudian psychoanalysis divided the psyche between the unconscious, which was grounded in biology, and the 'faculty of autonomous reason, lodged in consciousness and free of natural causes.' This division of ego from ID, although it rejected the Kantian formulation of rationality, was built on the same foundations, assuming a relationship between the mind and nature which needed explanation to preserve autonomy. Freud's conception intended to free the autonomous conscious from the determined subconscious, adding detail to the Enlightenment idea of human reason overcoming animal drives. (Tauber, 2009 pp.1-2)

⁹³As summarised in Courtois, 2006

⁹⁴Habermas, 2003 pp.57, as quoted in Courtois, 2006 pp.445-446

⁹⁵"In a tantalizing way many individuals have experienced just enough of creative living to recognize that for most of the time they are living uncreatively, as if caught up in the creativity of someone else, or of a machine." (Winnicott, 1971 pp.65, as quoted in Ffytche, 2007 pp.39)

⁹⁶And one avenue that is open to the transhumanist, wishing to address Habermas' concern about sincerity, might be to differentiate between paternalist predictions employed before birth (affecting future persons) and paternalist interventions used to predict and guide the action of current persons.

⁹⁷Fully autonomous individuals can be irrational, and irrationality can sometimes be an important part of decision-making and everyday interactions, particularly where not all the relevant information for rationality is available. Bortolotti 2020 pp. 52 Filling in missing information, for example, can be beneficial socially and can allow decisions in situations where reason alone would not allow it or would leave us 'dumbfounded'. Bortolotti 2018 pp. 247.

⁹⁸Radoilska 2012, for example describes a possible account of weakness of will as a secondary failure of agency to overcome a more fundamental 'akrasia', an inner conflict which is 'poorly resolved'.

⁹⁹US Agency for International Development statement – Available at : <https://www.usaid.gov/news-information/press-releases/sep-17-2021-executive-order-crisis-ethiopia>

¹⁰⁰BBC Online 7th July 2020 'Coronavirus: Trump moves to pull US out of World Health Organization' though this decision was, to some extent, neutralized by the results of the 2021 US Elections, and may be fully reversed by the new administration.

¹⁰¹The coordination, division of labour, and expertise needed to enact health policy in different global locations, where cultural norms and practices relevant to healthcare vary, and where identical resources are not consistently available.

¹⁰²GMC guidance to UK doctors. Nov 2020.

¹⁰³Charland 1998

¹⁰⁴In some cases, dignity after death is considered alongside respect for patient autonomy (or assumed wishes) after death as well as respect for living next of kin. See for example, cases of posthumous sperm-retrieval (Orr and Siegler, 2002; Hodson and Parker, 2019)

¹⁰⁵For example, the benefit of a donor organ for prospective recipients, the benefit to family of genetic testing, the legal benefit of a post-mortem examination, the consequences of posthumous sperm removal. Orr and Siegler 2002; Hodson and Parker 2019.

¹⁰⁶WHO, Research Ethics Review Committee (ERC) Consent forms Available at: <https://www.who.int/groups/research-ethics-review-committee/guidelines-on-submitting-research-proposals-for-ethics-review/templates-for-informed-consent-forms>

¹⁰⁷Though further consultations and second opinions are usually possible and are often encouraged.

¹⁰⁸As an example, there seems to be a difference between a case where a clinician is conversationally clumsy or heavy-handed in delivering the treatment option, to a patient who is distressed but otherwise able to decide, and cases of patients more properly described as vulnerable adults, receiving treatment without fully understanding, and being advised rather than informed as a result.

¹⁰⁹Personal implications of complex treatment options need not be technical, but might include recovery time, changes of full recovery, likelihood of scarring, necessary changes to diet and

lifestyle, impact on ability to work, reproductive status, chance of reoccurrence of condition, or impact on probable quality adjusted life years (QALY).

¹¹⁰Solomon 2015 describes the attractiveness of personalised medicine in many cases as something more like a commitment to patients' personhood than truly tailored medicine to their genetics, for example. pp. 228

¹¹¹Charland 1998

¹¹²The case caused scandal at the time (1964) and was reported initially by the New York World Telegram 20th January 1964. The news headlines also focused on the 'live cancer cells' first, and the deception second.

¹¹³Tversky and Kahneman 1981

¹¹⁴See Geurts 2013 for analysis of the results of this study and its implications.

¹¹⁵I will discuss the important role of affective, emotional, cognition in making rational decisions in the third dialogue. For more on the importance of affective responses in medical decisions see, for example, Charland, 1998

¹¹⁶Principle One, from the current GMC guidance to UK doctors. Nov. 2020.

¹¹⁷Principle Four, from the current GMC guidance to UK doctors. Nov. 2020.

¹¹⁸Provided the approach of the physician does not breach the guidance, individual approaches to this task fall to the professional judgement of individual physicians, as they judge appropriate, and are not prescribed to physicians. GMC guidance to UK doctors. Nov 2020.

¹¹⁹GMC guidance to UK doctors. Nov. 2020

¹²⁰Jeffrey 2016 discussed the phenomenon the NHS refers to as 'existential neglect' of patients. pp. 446

¹²¹The available WHO Consent form template uses malaria trials as an example, and clarifies that clinicians should ensure participants know they have not been selected for the trial because they are known to be particularly unwell, for example.

¹²²General optimism biases and success biases (people routinely place themselves in the higher percentiles for *i.e.*, *school test scores*) also affect beliefs like 'It is more likely that I am in the group receiving the real drug, not the placebo', in a randomly assigned control condition where groups are assigned by chance.

¹²³Organ donor rates are one such social limitation which can place restrictions on theoretically common available services and create long waiting lists in certain locations. Sex reassignment surgery, by contrast, has been known to experience restriction and long waiting lists because of the numbers of qualified specialists available.

¹²⁴Gigerenzer 2015

¹²⁵Mill 2001 (1859) pp. 13

¹²⁶As in Habermas' argument against genetic engineering for enhancement purposes, summarised in Courtois, 2006 pp. 447, "When the parents' preferences are genetically imposed on their offspring, the offspring are manipulated as objects rather than treated as the autonomous and free persons we normally meet in a communicative relationship." The solution to this problem proposed by some, including Habermas, is that only therapeutic uses of engineering (those used to cure disease) are acceptable, and that positive enhancement is unacceptable. As will be discussed in the second dialogue, the distinction between therapy and enhancement is hard to find, and the philosophical distinctions made also tend to point to a significant social and political role in determining what is acceptable, even given Habermas' solution.

Notes to Chapter 3: Second Dialogue

¹²⁷Solomon 2015 pp.6

¹²⁸Solomon 2015 pp.6

¹²⁹Solomon 2007 pp. 169

¹³⁰However, Solomon also notes, evidence-based approaches may help individual researchers avoid various kinds of personal bias, but they are not always as reliable as expected.

¹³¹ This is to identify as separate but not exclusively so, the interface consensus (between experts already in general agreement and society) from a technical academic consensus (where experts discuss opposing scientific positions until a consensus is reached).

¹³²Solomon 2015 pp. 229

¹³³Importantly, however, the different methods do not neatly slot together at different stages of medical research, evaluation, and implementation. There is a, sometimes messy, overlap and, Solomon suggests, ‘no hierarchy of methods’ 2015 pp. 229

¹³⁴Solomon 2015 pp. 169

¹³⁵Bostrom 2005a

¹³⁶In fact, Kuhn “rejected the traditional correspondence theory of truth and the related idea of cumulative progress toward a representational truth waiting out there for science to find it” (Nickles, 2017) and as a result his early critics took Kuhn to be anti-science because of this divergence. Kuhn’s criticism of the Enlightenment idea of cumulative progress, in questioning the infallibility of one universally successful ‘scientific method’ as the reason for scientific progress, left a gap. (For a revised account of the case against scientific method, drawing on Kuhn and Feyerabend, see for example Alan Chalmers’ chapter in Schuster, 1986)

¹³⁷Kuhn 1962 pp. 10

¹³⁸For a fuller explanation of the dangers of the public perception that there is no ‘rigorous scientific method’ that remains fixed, see for example Theocharis & Psimopoulos, 1987.

¹³⁹Bird 2007

¹⁴⁰Kastrup 2018 pp. 43

¹⁴¹Bostrom 2011

¹⁴²Kuhn 1962 pp. 118

¹⁴³About the particular aims of transhumanism, relating to these core capacities, see the primer chapter.

¹⁴⁴A point that has often been made about modernity in general as well as about transhumanism (for transhumanism in particular see, for instance, Church, 2019), is that the average person now lives well above the standard that would have been available to ancient kings. This much seems true, as our homes, social organisation, hobbies, our work, dress, and communication might make us alien to our ancestors. On the other hand, in physical appearance and core capacities, supposing a 21st century human child could be somehow raised by medieval royalty, we are nowhere near as jarringly unrecognisable or alien.

¹⁴⁵One such successful rodent trial was reported by O’isín Biotechnologies (Lewis, 2019; also, Baker et al. 2016)

¹⁴⁶Bostrom, 2011 makes this argument about the transhumanist goal of a more efficient epistemic community, and less general ignorance overall.

¹⁴⁷Skinner 1971 p10, emphasis added

¹⁴⁸Balaguer 2009

¹⁴⁹Compelling, despite reasonable criticism that Skinner’s particular “speculations are devoid of scientific content and do not even hint at general outlines of a possible science of human behavior.” In his review of ‘Beyond Freedom and Dignity’ Chomsky’s criticism of premature and overblown leaps made, based on flimsy science, to serve social objectives, are likewise, damning. “Skinner imposes certain arbitrary limitations on scientific research which virtually guarantee continued failure”.

¹⁵⁰A theory based on empirical observation, rules, and logical theory, as described by Danto, 1985 quoted in Reisch, 1991: “[t]he philosophical task was to provide a formal reconstruction of the language of science, conceived of as a logical edifice, resting upon observational reports, overarched by sentences of increasing generality and scope” p265

¹⁵¹A prominent current example can be found in companies like *23andme* which encourage genetic testing both to discover genetic heritage and to foresee illness with genetic markers. Although individuals sign consent forms allowing such companies to use their data, ethical questions have been raised about whether individuals understand the uses of such data well enough to consent, or whether the rate at which possible uses and analyses of the data changes should be of concern to those who have already given blanket consent. As an example, we can imagine an individual who was perfectly happy for companies to test for genetic variations relating to their own risk of illness, but not happy for any suggestive data about whether a person is more likely to have violent tendencies to be passed to the authorities. This is not an example geneticists would find concerning, and the science suggests that such analyses would be too speculative and tied to environmental factors to be worth pursuing, but i) individuals would find this kind of change concerning and ii) the technology might advance, within an individual’s lifetime, such that these analyses become precise enough to be worthwhile.

¹⁵²EAST, from the Behavioural Insights Team

¹⁵³Sunstein 2019

Notes to Chapter 3: Third Dialogue

¹⁵⁴Leong et al 2016

¹⁵⁵Bourgeron 2021 pp. 6

¹⁵⁶Wellcome Global Monitor 2019

¹⁵⁷In this case, an incident involving the mishandling of contaminated blood for transplant, resulting in illness, had been fresh in the public consciousness and blamed on the government. REF

¹⁵⁸See, for example, Giubilini, 2019

¹⁵⁹When automatic retrieval of corneas during autopsy was carried out in Georgia, the number of available corneas for transplant did increase fourfold, between 1978 and 1984. However, this practice raises concerns about the beliefs, preferences, and expectations of families about the dignified treatment of their loved ones during autopsy and burial, let alone the general belief that people should be able to decide what happens to their bodies after death, within reason, and to have those wishes respected where possible. Thaler and Sunstein, 2009 pp. 186

¹⁶⁰Tversky and Kahneman 1981

¹⁶¹See for example Kahneman 2011

¹⁶²The Behavioural Insights Team's EAST report identifies the difference between stated preferences and apparent preferences based on behaviour, and the literature on unconscious bias suggests that people are able to give reasons unrelated to the bias for their actions, under test conditions where a bias is in evidence.

¹⁶³Kreitchmann et al. 2019 for example, discuss the need to offset reporting biases because participants in studies are more likely to give answers they hope will be acceptable to other participants or to the researchers than to report accurately on their attitudes or reasons. This is not necessarily always consciously done.

¹⁶⁴In some studies participants did unwittingly avoid a stimulus as the result of bias or conditioning, while reasoning and reporting that they felt none of the fear, distaste, or trepidation toward the stimulus that their behaviour suggested. It seems likely, therefore, that experience (and awareness) of emotions might be reduced when behaviour is motivated by unconscious bias, though emotions themselves are cognitive.

¹⁶⁵The cognitive view of emotions as essential for reason, and particularly as essential for moral knowledge, is supported by studies, for example Damasio 1994.

¹⁶⁶Roeser 2006 pp. 693

¹⁶⁷Roeser 2006 pp. 692

¹⁶⁸Roeser 2010 pp. 240

¹⁶⁹Frank 2011; 1988

¹⁷⁰Roeser 2010 pp. 242

¹⁷¹Roeser 2010 pp. 242

¹⁷²James 1896

¹⁷³James 1896

¹⁷⁴James' example slightly predates the kind of global communication that makes more borderless religious community and shared credibility of traditions slightly more feasible, and perhaps the fact that his example no longer seems entirely fit, in a world where westerners might find something scintillating and electric in the prospect of an eastern religion, might serve as another example of how hypotheses can either match with, or jar against, an agent's experience of the world dependent on time and place. The apparent artifice of choosing traditions and practices from other nations than one's own, which made James' example nicely illustrative in 1896, is no longer culturally jarring in the same way, for many in 2020.

¹⁷⁵Trivial options are those where "the opportunity is not unique, when the stake is insignificant, or when the decision is reversible if it later prove unwise." James 1896

Notes to Chapter 4

¹⁷⁶For example, CRISPR gene editing, as described in the primer chapter.

¹⁷⁷Anomaly 2011

¹⁷⁸For an explanation of why enhancements should be even further beyond public health goods than tackling obesity or the causes of gunshot wounds, see earlier discussion of ‘treatment versus enhancement’ arguments about the limits of medicalization, that would exclude enhancement from medicine.

¹⁷⁹For a wider definition of public health, see for example Dees 2018

¹⁸⁰Even though ‘preventable deaths’ are really delayed rather than prevented, as far as the individual is concerned.

¹⁸¹Murray and Lopez 1996.

¹⁸²As in the revised DALY model, Murray and Lopez, 2013; Salomon et al., 2012

¹⁸³For example, from Gigerenzer, 2015; Miller and Gelinas 2013

¹⁸⁴By Kuhn, Müller, Heidbrink, Buyx, 2020

¹⁸⁵Thaler and Sunstein 2008

¹⁸⁶As in Anomaly, 2011 pp.252

¹⁸⁷Dworkin 1972 pp.69 as in Anomaly 2011 pp.252

¹⁸⁸Anomaly, 2011 pp.253

¹⁸⁹For example, Faden and Powers, as quoted in Anomaly 2011

¹⁹⁰Solomon, 2015 pp.210

¹⁹¹Solomon, 2015 pp.211

¹⁹²For example, Baum et al, 2005 as described in Solomon, 2015 pp.212

¹⁹³Solomon, 2015 pp.212

¹⁹⁴The literature describes the absolute risk difference, between individuals with a cancer risk of 4.95 per cent given routine screenings, and 5 per cent without, as comparatively insignificant.

¹⁹⁵Solomon describes the view as one developed in the 1960s which has remained popular, despite some evidence that routine screenings may not be of overall benefit to most patients. Solomon points out that, although there is not conclusive evidence that early detection saves lives in the way that was believed initially, there is also not conclusive evidence that early detection does not save lives, nor conclusively that it costs lives. (2015, pp.211)

¹⁹⁶Hughes 2002.

¹⁹⁷The European Commission’s compiling agency: Eurostat, the statistical office of the European Union Glossary 2021

¹⁹⁸Solberg et al, 2020 pp. 260

¹⁹⁹See, for example, Hughes 2005, Herlitz 2017

²⁰⁰The World Bank Disease Control Priorities study, the Global Burden of Disease study, and the US Panel on Cost-Effectiveness in Health and Medicine recommended a 3% time-discounting measure, to weigh more immediate costs and health outcomes more heavily than those in future years. This was a reduction from the 5% that had previously been used to weigh current cost and benefit as more important.

²⁰¹Anomaly 2011, Horne 2019

²⁰²Solomon 2015 pp. 218 explains changing guidelines in the advice on mammography screenings, despite missing consensus.

²⁰³Murray and Acharya 1997

²⁰⁴For example, from the Office for National Statistics in the UK (who publishes in 2018 an analysis of the trends in life expectancy since 1981, and the way life expectancy after 2018 would be calculated to adjust for a noticeable slowing effect since 2011), and the World Health Organisation (the last published global average life and health expectancy are from 2019. The global average life expectancy in 2019 was 73 years at birth).

²⁰⁵As in the case of mental health classifications in updated versions of the DSM, for example, which can have political, social and medical consequences for individuals seeking professional attention.

²⁰⁶Gigerenzer 2015

²⁰⁷Hughes 2005 pp.298-299

²⁰⁸Solomon 2015 pp. 224

²⁰⁹As in Solberg et al. 2020 pp.260

²¹⁰Murray and Lopez 2013

²¹¹de Grey et al. (2004) identify whole-body interdiction of lengthening of telomeres (WILTY), a negligible senescence strategy to undo the effects of bodily ageing, as the only realistic strategy to address the increased risk of cancer that faces human beings with lifespans as long as ours now tend to be. Globally and on average in global populations, risk of mortality from cancers of 36 types is

around 1 in 10, according to the most recent global study of cancer morbidity and mortality. Overall risk of morbidity from the 36 cancers is around 1 in 5 globally, making cancers one of the leading causes of global disease burden in most countries. (Global Cancer Statistics 2020, Sung et al. 2021 pp.219)

²¹²Report from the National Advisory Council on Bioethics 2015

²¹³Saghai 2013, pp.491

²¹⁴'MINDSPACE' from the Behavioural Insights Team. See Appendix 1.8 for a glossary of their behavioural insights tools.

²¹⁵MINDSPACE 2010 pp. 14

²¹⁶In Kahneman's *Thinking Fast and Slow* and other reports referenced by MINDSPACE to support intervention, reflective processes are 'SYSTEM 1' and automatic processes are 'SYSTEM 2'. The different systems each have roles in different parts of decision making and behaviour, are engaged at different times for different individuals, and can be prompted to switch between systems.

²¹⁷Nussbaum 2013, Nussbaum 2004

²¹⁸Nussbaum 2013

²¹⁹Nussbaum proposes a better appreciation for the pluralism of people's consciences, to help balance the inclusion of *i.e.*, religious beliefs with state authority. 2013 pp. 108

²²⁰It is important to note the meta-analysis does also suggest that availability and familiarity of technologies in any given country are a likely contributing factor that would tend to occur alongside education and the confidence of educators, and might also be responsible for some of the results seen. Weil and Rosen 1995 pp. 111

²²¹Weil and Rosen 1995 pp.129

²²²Council of Europe 1989 pp.1

²²³Weil and Rosen 1995, pp.96

²²⁴As used in Jay 1981

²²⁵And this concern is not very new either. See Kahane and Starr 1976 for a review of the impact of rapid social change on modes of education in the 1970s.

²²⁶Rosen 1987

²²⁷Agogo and Hess 2018

²²⁸Agogo and Hess 2018 pp. 588

²²⁹Agogo and Hess 2018 pp. 588

²³⁰Weil and Rosen 1995

²³¹For example, in Hofmann 2002

²³²Ménard 2010 pp. 236

²³³NHS Blood and Transplant 2021, which outlines the 2013-18 strategy

²³⁴As a very recent example, the approach taken in the 14th April WHO Covid-19 global strategy update.

²³⁵The Humanity+ manifesto and its advocacy for human development, as in the primer chapter.

²³⁶Gigerenzer 2015

²³⁷Menard 2010 on Thaler and Sunstein's rejection of financial incentives. 2010 pp. 233

²³⁸For example, Brown 2018 addresses the breadth of Stanley's definition of propaganda as 'the employment of a political ideal against itself', and distinguishes supporting propaganda (which can promote, for example, patriotic ideals) from undermining propaganda (which would, for example, present a notion of patriotism deliberately to undermine it or to repulse people from it). Brown 2018 pp. 196-197. Brown suggests that undermining propaganda does not clearly fit the more general definition of 'ideals employed against themselves' or at least that an undermining example of this more general propaganda is hard to find.

Notes to Chapter 5

²³⁹Hughes 2002

²⁴⁰As an interesting example, I consider the self-modification of artist Neil Harbisson, and the modification technology that allows wealthy individuals to insert smart technology to track pulse rate, under their skin. Guardian Online 2014

²⁴¹Pilsch 2017 pp. 4

²⁴²Dworkin 1988 pp. 9

-
- ²⁴³Though the use of ‘drugs’ as commonly meant is a more complex case because the sale and distribution of controlled substances is often regulated by law, even in countries where personal consumption is not strictly illegal.
- ²⁴⁴Though the consequences of attempted suicide and self-harm can be severe, and risk-taking behaviour which could cause harm to oneself or others is sometimes taken to be reason to limit the agency of individuals through *i.e.*, sectioning, conservatorships.
- ²⁴⁵Some organs like eyes, or tissue structures like faces retain peculiar status as taboo for transplant, even among individuals who are happy to donate their internal tissues. BBC News 2010
- ²⁴⁶Harris 2005, Brassington 2007, Chan and Harris 2009
- ²⁴⁷Hughes 2010
- ²⁴⁸Dove et al 2017 pp. 162
- ²⁴⁹Prainsack 2017 pp. 38
- ²⁵⁰Schleidgen et al 2015 pp. 20
- ²⁵¹Biomarkers, in this context, can be understood as molecular signs of illness or abnormal processes of the body (for example, observed in the blood) which can indicate or predict illness.
- ²⁵²Dove et al 2017 pp. 156
- ²⁵³Fischer et al 2015 pp. 25
- ²⁵⁴As in Arneson 2012 pp. 104, clinical interactions with patients are increasingly subject to bioethical developments which mean clinicians give more genetic advice to accompany more genetic tests.
- ²⁵⁵Dove et al 2017 pp. 156 - 157. Illustrative examples include cases where patients with heritable conditions refuse to inform family members who might be affected and refuse to allow clinicians to do so; cases where placental sampling would mean biological samples collected from one patient had the potential to help clinicians better understand how to treat another patient.
- ²⁵⁶WHO recommendations 2018 following Shenzhen
- ²⁵⁷Levin 2021 offers the most recent comprehensive and multi-disciplinary criticism of overzealous transhumanist ideology.
- ²⁵⁸Nuffield 2018 pp. viii
- ²⁵⁹The primer to transhumanism has more detail.
- ²⁶⁰Nuffield, 2018 pp. 100. The use of HGE on human subjects is currently illegal in the UK, as in other major global jurisdictions.
- ²⁶¹As found in the Transhumanist Manifesto (Humanity +)
- ²⁶²Courtois 2006 pp. 443
- ²⁶³Gyngell et al 2019 pp. 516
- ²⁶⁴Gyngell et al 2019 pp. 516
- ²⁶⁵Halpern 2019 pp. 2-3
- ²⁶⁶Halpern 2019 pp. 3-4
- ²⁶⁷Bostrom 2005a
- ²⁶⁸Hughes 2002
- ²⁶⁹Notably, the Methuselah Foundation has received support in recent years from Peter Thiel (PayPal, Palantir) and Larry Ellison (Oracle Corporation). Other investment in negligible senescence and life extension has come from Larry Page and Sergey Brin (Google) and from Mark Zuckerberg (Facebook). Elon Musk has also been a supporter of transhumanist projects for several years but has tended to favour a more Kurzweil-esque approach, including brain-computer interfacing and digital brain upgrades (as well as interplanetary travel).
- ²⁷⁰ Particularly when financial support comes from private individuals, rather than regulated medical bodies.
- ²⁷¹Halpern 2019
- ²⁷²Consider Thaler and Sunstein’s school cafeteria example, where children have a choice of fries or carrot sticks, which could be placed strategically to make either the more salient and attractive choice to the children at the point they were queuing with their trays. 2008 pp. 2
- ²⁷³Mongin and Cozic 2018 pp. 108
- ²⁷⁴Halpern 2019 pp. 172
- ²⁷⁵Dear, Dutton and Fox 2019
- ²⁷⁶See Appendix Fig 1.5
- ²⁷⁷Thaler and Sunstein 2008

²⁷⁸Report from the Organ Donation Taskforce, 2008 pp. 2 The report suggests that a 50% increase in donation within five years should be possible, given their recommendations.

²⁷⁹Report from the Organ Donation Taskforce 2008 pp. 5. The report also notes that the seriousness of the national DBD shortage becomes even more clear, given the willingness of patients, clinicians, and commissioners to accept the risks associated with rising numbers of LD procedures.

²⁸⁰This overall figure of 40% refusal was higher in some individual demographics within the UK population, and the report also addresses concerns over fairness and just distribution of resources. Report from the Organ Donation Taskforce, 2008 pp. 49. A later awareness campaign undertaken by NHSBT targeted certain cultural and religious demographics within the UK to address nation-wide inequalities of this sort affecting donor numbers and patient demand for organs.

²⁸¹Report from the Organ Donation Taskforce 2008 pp. 6

²⁸²Of the 1% of deaths the NHS indicates happen in a way that makes transplant viable. From *'English donation law is changing- easy read' pamphlet. NHS Blood and Transplant*

²⁸³In a summary on the use of polling by health bodies like the NHS 'What do Opinion Polls tell us?' Ben Page's presentation at National Donation and Transplantation Congress, September 2013 used evidence drawn from IPSOS Mori data.

²⁸⁴Clinton and Grisson 2015

²⁸⁵Bishop 2005 pp. 23

²⁸⁶Clinton and Grisson 2015

²⁸⁷Consider, for example, this comment on the value of public opinion polling, from Roger Mortimore as Director of Political Analysis, for Ipsos MORI Almanac 2013 pp. 40. "In our business, we have to be very suspicious of 'interesting' results – public opinion is actually relatively predictable (even if the predictability is sometimes limited to knowing that the opinions themselves are volatile). Sudden or dramatic changes to the pattern of things, unless there is a very clear reason for them, are more likely to indicate that something has gone wrong with the measurement than that public opinion has really flipped. So, when we have measured attitudes to the same question regularly over forty years and got the same answer (within the margin of error), our natural reaction if we suddenly get a poll eight percentage points out of line is to scratch our heads and wonder what has gone wrong."

²⁸⁸The Behavioural Insights Team – also known as the Nudge Unit – is now a social purpose company. It is partly owned by the Cabinet Office, employees and Nesta. From the Behavioural Insights Team's 2010 report published by the Cabinet Office, 'MINDSPACE Influencing behaviour through public policy', nudges are subdivided into the following categories: Messenger, Incentive, Norms, Defaults, Salience, Priming, Affect, Commitments, Ego. As in chapter 3, incentives are not always accepted by nudge theorists as a non-coercive behavioural intervention, because of pre-existing social inequalities within the target population of the nudge that can make it more forceful for some than others.

²⁸⁹See Appendix Figs 1.2-1.6 for more examples of emotive posters from public health agencies in the UK, which rely on assumed commitments to others.

²⁹⁰ For another example of an informative nudge relevant to the Organ Donation case, see Appendix Fig 1.6 'O negative blood can be given to anyone'.

²⁹¹Both are examples of NHS England Public Health posters: 'Have you joined the Organ Donor register?' 2012 NHS UK leaflet; 'Drop a drink size' NHS UK in partnership with Public Health Liverpool and Fewer Units (active 2013-2015)

²⁹²As in Stanley 2015, Brown 2018

²⁹³Report from the Organ Donation Taskforce 2008

²⁹⁴EAST, Behavioural Insights Team Report 2014 pp. 43

²⁹⁵Gigerenzer 2015

²⁹⁶The Organ Donation Taskforce (2008) identified the need to work slowly, and to identify important target demographics over the course of their work, sensitively and with flexibility.

²⁹⁷ Schmidt 2017 pp. 405

²⁹⁸Gigerenzer 2015

²⁹⁹Mongin and Cozic 2018

³⁰⁰Similarly, in Dove et al 2017, the example of the patient who must decide whether their doctor is at liberty to disclose important information about the possible risk of illness or need for tests to another, biologically related, patient in their care. pp. 156-157

³⁰¹Report from the Organ Donation Taskforce 2008 pp. 9

³⁰²Gyngell et al 2019: Applications of HGE have been shown to reduce incidence of (for example) chronic polygenic diseases like coronary heart disease (27 associated genetic mutations) or coronary artery disease (35 known genes involved in risk) and include PGT/IVF screening of embryos. Selection based on screening (whole genome or whole exome sequencing) alone would only limit numbers of embryos, perhaps to figures below 1% for couples involved in IVF. Sequencing combined with targeted editing would make the number of viable embryos with low risk for chronic diseases much higher.

³⁰³Thaler and Sunstein, as in Saghai 2013 pp. 488

³⁰⁴For example, if I fail to brush my teeth tomorrow morning the ill-effects on my health will not be apparent in any serious way. However, if I fail habitually, I will eventually see the urgency and salience of good dental care, but by then it could be much too late to make the right choice.

³⁰⁵Petticrew et al (2020) describe sludge as an anti-nudge: sludge is the opposite of making something easy, attractive, social, or timely.

³⁰⁶Petticrew et al 2020

³⁰⁷Gyngell et al 2019 pp. 520

³⁰⁸Elliot 1989 pp. 163

³⁰⁹As Sunstein explains, the cascading effect of the most confident free actors moving first, followed by those who need to see one other person act first, followed by those who are moved only when a few seem to act together... etc. Sunstein 2019 pp. 26

³¹⁰Nuffield 2018 pp. 82 Explains the 'expressivist objection' to enhancement, that some interventions by their availability express hostile views towards individuals whose bodies do not fit the implied norm or expectation of medical treatment.

³¹¹Saul 2017 pp. 8

³¹²See, for example, Bortolotti and Nagasawa 2009, Fischer and Mitchell-Levin 2014,

³¹³Hughes 2002

³¹⁴Sunstein 2019

³¹⁵"Protection of biometric information of children in schools". Department for Education. 13 December 2012

³¹⁶The protections needed for the ethical collection of sensitive biometric data were not in place, and the use of biometrics for trivial items like library books seemed disproportionate, as biometrics were primarily used to keep prison records of attendance and activity, at the time.

³¹⁷Which was missing during the early years of the use of biometrics, with most parents unaware of its use in schools. Source: Al Jazeera 2019 All Hail the Algorithm, BBC News online 2007 Child Fingerprint Plan Considered

³¹⁸At which point over 2 million school children were estimated to have been fingerprinted for the scheme. Al Jazeera 2019, Guardian Online 2006

³¹⁹Sprokkereef and De Hert 2007

³²⁰Agogo and Hess 2018

³²¹Agogo and Hess 2018 pp. 587

³²²Davis and Davidson 2006 pp. 31

³²³Davis and Davidson explain that the law in England was more complex and more strictly defined, criminalising both the intent to provide and the intent to seek, and including exemption clauses for certain likely consequences of pregnancy for the pregnant woman. Because Scottish law was less detailed, the shift of responsibility for determining when termination of pregnancy was essential to safeguard the patient (the pregnant individual) so the judgement could be more easily handed to clinicians as a purely medical decision about treatment options. 2006 pp. 31

³²⁴Davis and Davidson 2006 pp. 31

³²⁵Davis and Davidson 2006 footnote 31 on pp. 35

³²⁶Davis and Davidson 2006 pp. 34

³²⁷Davis and Davidson also highlight the significance of religious views in repressing the open discussion or enthusiastic delivery of medical terminations to women in Glasgow during this time, contrasting the city with Aberdeen. 2006 pp. 37

³²⁸Baird, as quoted in Davis and Davidson 2006 pp. 34

³²⁹Russo 2014 pp. 240

³³⁰Charland 1998 pp. 68

³³¹Charland 1998 pp. 71

³³²Russo 2014 pp. 241

-
- ³³³Guardian online 2019 quotes the American Civil Liberties Union (ACLU) who opposed the legislation in Kentucky.
- ³³⁴Nussbaum 2013 pp. 390
- ³³⁵Though Thaler and Sunstein do not suggest choice, or the ‘maximisation’ of choice is a good thing. The more choices in the set, the more guidance is likely needed. 2008 pp. 94
- ³³⁶Nussbaum, 2015 pp.3
- ³³⁷Summarised in chapter 3
- ³³⁸Fischer and Mitchell-Yellin 2014
- ³³⁹Bortolotti and Nagasawa 2009
- ³⁴⁰Sparrow 2019
- ³⁴¹Fukuyama 2003
- ³⁴²Affective responses towards computers are described in much of the existing research as the literature is primarily about the introduction of personal computers in the 1970s.
- ³⁴³Fahlquist 2019 pp. 214
- ³⁴⁴As discussed in chapter 5
- ³⁴⁵Nussbaum 2013
- ³⁴⁶Nussbaum 2013 pp. 6
- ³⁴⁷Nussbaum 2013 pp. 132
- ³⁴⁸Nussbaum 2013
- ³⁴⁹Nussbaum 2018
- ³⁵⁰The suggestion of a shift of focus to national service as a unifying tribulation for young people might also present other, potentially distracting values, calling on different and potentially conflicting commitments as well as the wellbeing of others, relating to combat, warfare, and firearms training.
- ³⁵¹Pilsch in his 2017 explanation of the pursuit of transhumanism described it as a situation wherein “[n]ature is no longer driving our species’ evolutionary bus.” (2017 pp. 9) Following the Shenzhen trial in 2018, more global stakeholders have taken seriously the challenge that a new driver should be appointed soon.
- ³⁵²McStay 2020 describes a ‘weak consensus’ about privacy of data relating to the use of emotional AI surveillance and identifies an opportunity for the apparent consensus to be put into action to guide policy work. The consensus he describes is a general preference for privacy, with ‘over half of UK citizens are ‘not OK’ with the principle of emotion detection’ pp. 10, and that the gap in European law on emotional detection AI should be addressed.
- ³⁵³McConnell and Kennett 2006 pp. 63
- ³⁵⁴Kass 1997 pp. 18
- ³⁵⁵McConnell and Kennett 2016 pp. 63-64
- ³⁵⁶People are more likely to describe the character of their interlocutor as ‘warm and friendly’ if they are holding a warm mug during a conversation (Williams and Bargh 2008), people are more likely to condemn a moral failing harshly if they are seated at an untidy desk at the time of the judgement, and so on (Schnall et al 2008).
- ³⁵⁷Damasio 2013 pp. 143, McConnell and Kennett 2016 pp. 65 – 66, Sapolsky 2017 pp.
- ³⁵⁸Ritov and Baron 2010 pp. 659
- ³⁵⁹Bostrom and Ord 2006
- ³⁶⁰Wells 2008
- ³⁶¹Brownwell 2019 quotes Noah Davidsohn on his recent research.
- ³⁶²For example, ‘CYBORG CRAZE SEES MORE THAN 4,000 SWEDES INSERT CHIPS UNDER THEIR SKIN’ Cuthbertson for the Independent Online, 26th October 2018, and ‘Pig brains kept alive outside body for hours after death’ Reardon for Nature Online 17th April 2019, as a common archetype of body shock reporting in press reporting on both self-enhancement of the body and animal trials of enhancement.
- ³⁶³Sunstein 2019
- ³⁶⁴Hay 2021
- ³⁶⁵Analysis of the influence of tensions over the incorporation of birth control into general practice by Davis and Davidson (2006), and Hay (2021) show that already vulnerable women and those with already limited access to healthcare in small communities were most affected by social tensions surrounding the inclusion of abortion into the professional domain of medicine in the 1960s.

Notes to Conclusion and Postscript

³⁶⁶James, 1896, II

³⁶⁷James, 1896, III

³⁶⁸James, 1896, III

³⁶⁹As in the interpersonal autonomy of interested groups, described in Dove et al. 2017

³⁷⁰What's so bad about feeling good? A 1968 film, directed by George Seaton, and starring George Peppard and Mary Tyler Moore.

³⁷¹Dr David King established 'Independent SAGE' along with a panel of other experts in the medical and scientific community, following concern about the advice and the use of advice by the UK government from their SAGE group. One of the values expressed by Independent SAGE was the belief that 'the public wants to hear about the science surrounding COVID-19 and the thinking that underpins the government's strategy.' [From What is Independent Sage?, Available Online]

³⁷²SAGE – the Scientific Advisory Group for Emergencies

³⁷³The first report of a successful vaccine was released during November 2020.

³⁷⁴For example, the UK Government's 'Protect the NHS' public health campaign message 2020.

Bibliography

ADDISON, C. and TAYLOR-ALEXANDER, S., 2015. Gene Editing and Germ-line Intervention: The Need for Novel Responses to Novel Technologies. *Molecular Therapy*, 23(11), pp. 1678-1680.

AGGARWAL, A., DAVIES, J. and SULLIVAN, R., 2014. "Nudge" in the clinical consultation- an acceptable form of medical paternalism? *BMC Medical Ethics*, 15, pp. 31.

ALFANO, M. and ROBICHAUD, P., forthcoming. Nudges and other moral technologies in the context of power: Assigning and accepting responsibility. In: D. BOONIN, ed, *Handbook of Philosophy and Public Policy*. Palgrave.

ALFANO, M., 2017. Twenty-first century perspectivism: The role of emotions in scientific inquiry. *Studi di Estetica*, 7(1), pp. 65-79.

ALFANO, M., 2020. Can Real Social Epistemic Networks Deliver the Wisdom of Crowds? In: T. LOMBROZO, J. KNOBE and S. NICHOLS, eds, *Oxford Studies in Experimental Philosophy* (Vol. 3). Oxford University Press.

ALFANO, M., CARTER, J. and CHEONG, M., 2018. Technological Seduction and Self-Radicalization. *Journal of the American Philosophical Association*, 4(3), pp. 298-322.

ANITEYE, P., O'BRIEN, B., MAYHEW, S.H., 2016 Stigmatized by association: challenges for abortion service providers in Ghana. *BMC Health Services Research* (2016) Published online September 2016 16:486

ANOMALY, J 2011 Public Health and Public Goods. *Public Health Ethics* Vol.4 No.3 pp. 251-259

ARCHER, A., 2018. Are We Obligated to Enhance for Moral Perfection? *Journal of Medicine and Philosophy*, 43, pp. 490-505.

-
- ARENDRT, H., 1998 *The Human Condition* (1958) The University of Chicago Press, Ltd. London
- AYER, A., 1954. *Freedom and Necessity*. Philosophical Essays. Reprinted with permission from Proceedings of the British Academy from 1948. 1962 edn. The British Academy, pp. 1-25.
- BALAGUER, M., 2009 *Synthese* (2009 Why there are no good arguments for any interesting version of determinism 168: pp. 1–21
- BALL, W. and HOLLAND, S., 2009 *The Fear of New Technology: A Naturally Occurring Phenomenon*. *The American Journal of Bioethics* 9:1 pp. 14-16
- BARBALET, J., *Boredom and Social Meaning*. *British Journal of Sociology*. Vol. 50 No. 4 pp. 631-646
- BAVELIER, D., SAVULESCU, J., FRIED, L., FRIEDMANN., T., LATHAN, C.E., SCHÜRLE, S., and BEARD, JR., 2019. *Rethinking Human Enhancement as Collective Welfarism*. *Nature Human Behaviour* 2019 March; 3(3): pp. 204–206.
- BAYLIS, F., DARNOVSKY, M., HASSON, K., and KRAHN, T.M., 2020 *Human Germline and Heritable Genome Editing: The Global Policy Landscape*. *The CRISPR Journal*. Vol.3 No.5 pp. 365-377
- BBC News Online 7th July 2020 *Coronavirus: Trumo moves to pull US out of World Health Organisation*. Available at: <https://www.bbc.co.uk/news/world-us-canada-53327906> [Accessed September 2021]
- BECKER, E., and RINCK, M., 2004 *Sensitivity and response bias in fear of spiders*, *Cognition and Emotion*, 18:7, pp. 961-976
- Behavioural Insights Team 2014 *EAST Four simple ways to apply behavioural insights 2014* In partnership with the Cabinet office. Available at: https://www.bi.team/wp-content/uploads/2015/07/BIT-Publication-EAST_FA_WEB.pdf [Accessed August 2019]
- BESKOW, L. M., HAMMACK-AVIRAN, C. M., BRELSFORD, K.M., 2020 *Thought Leader Comparisons of Risks in Precision Medicine Research Ethics and Human Research* 42, no. 6 (2020) pp. 35-40.
- BINGHAM, P., VERLANDER, N. and CHEAL, M., 2004. *John Snow, William Farr and the 1849 outbreak of cholera that affected London: a reworking of the data highlights the importance of the water supply*. *Public Health*, 118, pp. 387-394.
- BISHOP, G.F., 2005 *The Illusion of Public Opinion: Fact and Artefact in American Opinion Polls*. Rowman & Littlefield.
- BORTOLOTTI, L. and ANTROBUS, M., 2015. *Costs and benefits of realism and optimism*. *Current Opinion in Psychiatry*, 28(2), pp. 194-219.
- BORTOLOTTI, L., 2018. *Optimism, Agency, and Success*. *Ethical Theory Moral Practice*, 21, pp. 521-535.
- BORTOLOTTI, L., 2020 *The Epistemic Innocence of Irrational Beliefs*. Published to Oxford Scholarship Online: August 2020

-
- BOSTROM, N., 2003. Human Genetic Enhancements: A Transhumanist Perspective. *Journal of Value Inquiry*, 37(4), pp. 493-506.
- BOSTROM, N., 2005 *Transhumanist Values in Ethical Issues for the 21st Century*, ed. Frederick Adams (Philosophical Documentation Center Press, 2003); reprinted in *Review of Contemporary Philosophy*, Vol. 4, May 2005.
- BOSTROM, N., 2005a The Fable of the Dragon Tyrant. *Journal of Medical Ethics*, 31(5).
- BOSTROM, N., 2008. Why I Want to be a Posthuman When I Grow Up. In: B. GORDIJN and R. CHADWICK, eds, *Medical Enhancement and Posthumanity*. Springer, pp. 107-137.
- BOSTROM, N., in L. Zonneveld, H. Dijkstra and D. Ringoir (eds.) 2008. *Reshaping the Human Condition: Exploring Human Enhancement* The Hague: Rathenau Institute, in collaboration with the British Embassy, Science and Innovation Network and the Parliamentary Office of Science & Technology. pp. 29-36
- BOSTROM, N. and ORD, T., 2006. The Reversal Test: Eliminating Status Quo Bias in Applied Ethics. *Ethics*, 116(4), pp. 656-679.
- BOSTROM, N. and ROACHE, R., 2010. Smart Policy: Cognitive Enhancement and the Public Interest. *Contemporary Readings in Law and Social Justice*, 2(1), pp. 68-84.
- BRADY, M., 2010 *Virtue, Emotion and Attention* *Metaphilosophy*, Vol. 41, No. 1/2, SPECIAL ISSUE: VIRTUE AND VICE, MORAL AND EPISTEMIC (January 2010), pp. 115-131
- BRATMAN, M., 1979 Practical Reasoning and Weakness of the Will *Noûs*, 13: pp. 131–51
- BROSNAN, S. and DE WAAL, F., 2003. Monkeys reject unequal pay. *Nature*, 425, pp. 297-299.
- BROWN, E., 2018 Propaganda, Misinformation, and the Epistemic Value of Democracy. *Critical Review, A Journal of Politics and Society*. 30:3-4, pp. 194-218
- BROWNWELL, L., 2019 *Wyss Institute for Biologically Inspired Engineering at Harvard University: Combination gene therapy treats multiple age-related diseases in mice*. November 4th 2019. Available at: <https://wyss.harvard.edu/news/combination-gene-therapy-treats-multiple-age-related-diseases/> [Accessed September 2021]
- BUSS, A., 1997, "Weakness of Will," *Pacific Philosophical Quarterly*, 78(1): 13–44.
- BUTTONWOOD., 2012. In praise of pessimists. *The Economist*, (January 28 2012 edition.), pp. 71.
- Cabinet Office Behavioural Insights Team 2010 *Applying behavioural insight to health*. December 2010. United Kingdom. Available online: <https://www.bi.team/publications/mindspace/> [Accessed September 2021]
- CALVERT, J., 2014. Governing in the context of uncertainty *Hastings Center Report*, 44(6), pp. S31-S33.
- CAREL, H., 2016. *Phenomenology of Illness*. Oxford University Press. Published to Oxford Scholarship Online: September 2016

CAVALIERE, G., 2019. WHO Expert Advisory Committee on Developing Global Standards for Governance and Oversight of Human Genome Editing Background Paper. The Ethics of Human Genome Editing. Commissioned by the World Health Organization to serve as a background document for the first meeting of Expert Advisory Committee on Developing Global Standards for Governance and Oversight of Human Genome Editing (18-19 March 2019).

CHAIB, F., 2019 WHO launches global registry on human genome editing. 29 Aug 2019 news release. Available at: <https://www.who.int/news/item/29-08-2019-who-launches-global-registry-on-human-genome-editing> [Accessed August 2021]

CHARLAND, L.C., 1998 Is Mr. Spock Mentally Competent? Competence to Consent and Emotion. January 1998 *Philosophy, Psychiatry & Psychology* 5(1): pp. 67-81

CHIGNELL, A., 2018. The Ethics of Belief. The Stanford Encyclopedia of Philosophy. ZALTA, EN., ed, Spring 2018 edn. Available at: <https://plato.stanford.edu/archives/spr2018/entries/ethics-belief/>. [Accessed September 2021]

CHOMSKY, N., 1971. The Case Against BF Skinner. *New York Review of Books*, (December 30th Issue).

Church Lab 2021 Harvard Molecular Technologies. Last edited 11th August 2021 Available at: <https://arep.med.harvard.edu/> [Accessed September 2021]

CHURCH, G.M., 2017 'QnAs with Alta Charo and George Church' Interview by Prashant Nair. *Proceedings of the National Academy of Sciences of the United States of America*. June 6th 2017 Vol.114 No.23 pp. 5769–5771

CISLER, J.M., OLATUNJI, B.O., LOHR, J.M., and WILLIAMS, N.L., 2009 Attentional bias differences between fear and disgust: Implications for the role of disgust in disgust-related anxiety disorders. *Cognition and Emotion*. 23:4, pp. 675-687

CLARK, J., 2013. Integrating basic and higher-cognitive emotions within a common evolutionary framework: Lessons from the transformation of primate dominance into human pride. *Philosophical Psychology*, 26(3), pp. 437-460.

CLARKE, S., SAVULESCU, J., COADY, T., GIUBILINI, A., and SAGAL, S., 2016 The Ethics of Human Enhancement: Understanding the Debate. Published to Oxford Scholarship Online: November 2016

CLIFFORD, W., 1999. The ethics of belief, 1877. In: T. MADIGAN, ed, *The ethics of belief and other essays*. Amherst, MA: Prometheus, pp. 70-96.

CLINTON, J.D., and GRISSOM, J.A. 2015 Public information, public learning and public opinion: democratic accountability in education policy. *Journal of Public Policy*. July 2015, pp. 1-31

COHEN, J., 2019, Aug. 1, 2019, 11:30 AM. The untold story of the 'circle of trust' behind the world's first gene-edited babies. *Science Asia/Pacific Biology CRISPR in China*.

COIN, A., and DUBLJEVIC, V., 2020 The Authenticity of Machine-Augmented Human Intelligence: Therapy, Enhancement, and the Extended Mind. *Neuroethics* 19th October 2020
194

Council of Europe 1989 The information society -- A challenge for education policies? Five cases studies: Denmark, France, Italy, Spain, Turkey. (ERIC Document Reproduction Service No. ED327207, papers presented at the Standing Conference of European Ministers of Education, Istanbul, October 11-12, 1989)

COURTOIS, S., 2006 Genetic Engineering, Moral Autonomy, and Equal Treatment. *The Monist*, Vol. 89, No. 4, Genetics and Ethics October 2006 pp. 442-465

CUTHBERTSON, A., 2018. Swedish Cyborg Craze Sees More Than 4,000 Swedes Insert Chips Under Their Skin. *The Independent Online*. Available at:

<https://www.independent.co.uk/life-style/gadgets-and-tech/news/sweden-cyborg-rfid-chips-biohackers-biohax-a8601601.html> [Accessed September 2021]

CYRANOWSKI, D., 2018, 28 November 2018. CRISPR-baby scientist fails to satisfy critics. *Nature* (564), 13-14.

CYRANOWSKI, D., 2020, 03 January 2020. What CRISPR-baby prison sentences mean for research. *Nature News* (577), 154-155.

DAMASIO, A., and CARVALHO, G.B., 2013 The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews: Neuroscience*. Vol.14 February 2013 pp. 143-152

DAVIDSON, D., 1980, *Essays on Actions and Events*, Oxford: Clarendon Press

DAVIS, J., 2005. Life-extension and the Malthusian objection. *Journal of Medical Philosophy*, 30, pp. 27-44.

DAVIS, G., and DAVIDSON, R. 2006 "A Fifth Freedom" or "Hideous Atheistic Expediency"? The Medical Community and Abortion Law Reform in Scotland, c.1960–1975 *Medical History*, 2006, 50: pp. 29–48

DE GREY, A., 2005. Resistance to debate on how to postpone ageing is delaying progress and costing lives. *EMBO Reports*, (October 14).

DE GREY, A., 2007 Life Span Extension Research and Public Debate. *Studies in Ethics, Law, and Technology*, Vol. 1, Iss. 1 [2007], Art. 5 pp. 1-13

DE WAAL, F. and SHERBLUM, S., 2018. Bottom-up morality: The basis of human morality in our primate nature. *Journal of Moral Education*, 47(2), pp. 248-258.

DEAR, K., DUTTON, K., and FOX, E., 2019 Do 'watching eyes' influence antisocial behavior? A systematic review & meta-analysis. *Evolution and Human Behaviour* 40 (2019) pp. 269-280

DEEDAT, S., KENTEN, C., MORGAN, M., 2013 What are effective approaches to increasing rates of organ donor registration among ethnic minority populations: a systematic review *BMJ Open* 2013; 3:e003453 pp. 1-13

DEES, R., 2018 Public Health and Normative Public Goods. *Public Health Ethics*. Vol.11 No.1 2018 pp. 20-26

DOUGLAS, H., 2016 Values in Science. *The Oxford Handbook of Philosophy of Science* Ed. Paul Humphreys. Online Publication Date Dec 2015

-
- DOVE, E.S., KELLY, S.E., LUCIVERO, M.M, DHEENSA, S., PRAINSACK, B., 2017 Beyond individualism: Is there a place for relational autonomy in clinical practice and research? *Clinical Ethics* 2017 Vol.12(3) pp. 150-165
- DWORKIN, R., 1988 *The Theory and Practice of Autonomy*. Cambridge University Press
- ELLIOT, R., 1989 The Rights of Future People. *Journal of Applied Philosophy*, 1989, Vol.6 No.2 (1989) pp. 159-169
- ESTEP, P.W., and CHURCH, G.M., 2020 Transparency is key to ethical vaccine research. *Science* 18th December 2020 Vol.370 Issue 6523 pp. 1422-1423
- FAHLQUIST, J.N., 2019 Public Health and the Virtues of Responsibility, Compassion and Humility. *Public Health Ethics* Vol.12 No.3 2019 pp. 213-224
- FARMER, R. and SUNDBERG, N., 1986. Boredom Proneness – The Development and Correlates of a New Scale. *Journal of Personality Assessment*, 50(1), pp. 4-17.
- FFYTCHÉ, M. 2007 The Most Obscure Problem of All: Autonomy and its Vicissitudes in the interpretation of Dreams. In *PSYCHOANALYSIS AND HISTORY* (2007) 9(1)
- First Report of the Advisory Committee on Developing Global Standards for Governance and Oversight of Human Genome Editing. 2019. WORLD HEALTH ORGANISATION, Geneva, 18-19 March 2019.
- FRANK, R.H., 2011 The Strategic Role of the Emotions. *Emotion Review* Vol.3 No.3 July 2011 pp. 252–254
- FUKUYAMA, F., 2003 *Our Posthuman Future: Consequences of the Biotechnology Revolution*. London: Profile Books 2003
- GENDOLLA, G., 2017. Comment: Do Emotions Influence Action? – Of Course, They Are Hypo-Phenomena of Motivation. *Emotion Review*, 9(4), pp. 348-350.
- GIGERENZER, G., 2015 On the Supposed Evidence for Libertarian Paternalism. *Review of Philosophy and Psychology*. (2015) 6: pp. 361–383
- GIUBILINI, A., 2019 *The Ethics of Vaccination*. Palgrave Macmillan
- GORDON, A., TUSKEVICIUTE, R. and CHEN, C., 2013. A multimethod investigation of depressive symptoms, perceived understanding, and relationship quality. *Personal Relationships*, 20, pp. 635-654.
- GREELY, H.T., 2019 CRISPR'd babies: human germline genome editing in the 'He Jiankui affair'. 13th August 2019 *Journal of Law and the Biosciences*, pp. 111–183
- HAGMANN, D., HO, E. and LOEWENSTEIN, E., 2019. Nudging out support for a climate tax. *Nature*, 9(June 2019), pp. 484-489.
- HAKER, H., 2019. Habermas and the Question of Bioethics. *European Journal for Philosophy of Religion*, 11(4).
- HALPERN, D., 2019 *Inside the Nudge Unit: How small changes can make a big difference*. (with Owain Service and the Behavioural Insights Team) Penguin Random House UK.

-
- HANSEN, JESPERSEN, 2013 Nudge and the Manipulation of Choice A Framework for the Responsible Use of the Nudge Approach to Behaviour Change in Public Policy. *European Journal of Risk Regulation* Vol.4 Issue 1 pp. 3-28
- HARE, R., 1981. *Moral Thinking: Its Levels, Method, and Point* Published to Oxford Scholarship Online: November 2003. November 2003 edn.
- HARRIS, J., 2005 Scientific research is a moral duty. *Journal of Medical Ethics*, 31(4), pp. 242–248.
- HARRIS, J. 2011 Moral Enhancement and Freedom. *Bioethics* Volume 25 Number 2 2011 pp 102–111
- HAY, K., 2021 'More than a defence against bills': feminism and national identity in the Scottish abortion campaign, c. 1975–1990, *Women's History Review*, 30:4, pp. 594-612
- HERISSONE-KELLY, P., 2012 Habermas, Human Agency, and Human Genetic Enhancement The Grown, the Made, and Responsibility for Actions. *Cambridge Quarterly of Healthcare Ethics* (2012) 21: pp. 200-210
- HIGGINS, W.C., ROGERS, W.A., BALLANTYNE, A. and LIPWORTH, W., 2020. Against the Use and Publication of Contemporary Unethical Research: The Case of Chinese Transplant Research. *Journal of medical ethics*, 46, pp. 678-684.
- HIMMERICH, H., BENTLEY, J., KAN, C., and TREASURE, J., 2019 Genetic risk factors for eating disorders: an update and insights into pathophysiology. *Therapeutic Advances in Pharmacology* 2019 Vol.9: pp. 1-20
- HODSON, N., and PARKER, J., 2019 The ethical case for non-directed postmortem sperm donation. *Journal of Medical Ethics* 2020;0: pp. 1–4
- HOFMANN, B., 2002 Technological medicine and the autonomy of man. *Medicine, Health Care and Philosophy* 5: pp. 157–167
- HORNE, C.L. 2019 Public Health, Public Goods, and Market Failure *Public Health Ethics* Vol.12 No.3 2019 pp. 287-292
- HUANG, C. and BAUM, M., 2012. Nudge Ethics: Just a Game of Billiards? *The American Journal of Bioethics*, 12(2), pp. 22-24.
- HUGHES, J., 2005 Palliative Care and the QALY Problem *Health Care Analysis*, Vol. 13, No. 4, pp. 289 - 301
- HUGHES, J., 2010. Contradictions from the Enlightenment Roots of Transhumanism. *Journal for Medicine and Philosophy*, 35, pp. 622-640.
- HUME, D., 1994. Idea of a perfect commonwealth. In: K. HAAKONSSON, ed, *Hume: Political Essays*. Online publication date 2012 edn. *Cambridge Texts in the History of Political*
- LEONG, M., OLNICK, S., AKMAL, T., COPENHAVER, A., and RAZZAK, R., 2016 How Islam Influences End-of-Life Care: Education for Palliative Care Clinicians. *Journal of Pain and Symptom Management* Vol.52 No.6 December 2016 pp. 771-774

Is opposition to genetically modified food irrational? 3 June 2015, [Homepage of BBC News, Science and Environment] Available at: <https://www.bbc.co.uk/news/science-environment-32901834> [Accessed August 2019].

JAMES, W., 1884. What is an emotion? *Mind*, 9(34), pp. 188-205.

JAMES, W., 1896. *The Will to Believe. An Address to the Philosophical Clubs of Yale and Brown Universities.* The New World, June 1896, [Transcribed into HTML by William O'Meara, copyright, 1997].

JEFFREY, D., 2016 Empathy, sympathy and compassion in healthcare: Is there a problem? Is there a difference? Does it matter? *Journal of the Royal Society of Medicine*; 2016, Vol. 109(12) pp. 446–452

JUENGST, E.T., 1997 Can Enhancement Be Distinguished from Prevention in Genetic Medicine? *The Journal of Medicine and Philosophy* Apr;22(2): pp. 125-42

KAHANE, R., and STARR, L., 1976 The Impact of Rapid Social Change on Technological Education: An Israeli Example. *Comparative Education Review* Vol.20 No.2 June 1976 pp. 165

KAHNEMAN, D., 1977. Intuitive Prediction: biases and corrective procedures. Decision research. A branch of perceptrons. TECHNICAL REPORT PTR-1042-7746. Contract N00014-76-C-0074. Sponsored by DARPA.

KAMENS, S.R., 2011 On the Proposed Sexual and Gender Identity Diagnoses for DSM-5 History and Controversies *The Humanistic Psychologist*, 39: pp. 37–59

KASS, L., 1985 *Toward a More Natural Science: Biology and Human Affairs.* By Leon R. Kass. New York: Free Press

KASS, L., 2003 *Beyond Therapy: Biotechnology and the Pursuit of Human Improvement.* Presented at the President's Council on Bioethics 2003 meeting. Available at: <https://bioethicsarchive.georgetown.edu/pcbe/background/kasspaper.html> [Accessed September 2021]

KASTRUP, B., 2018 The Next Paradigm. *Future Human Image*, Volume 9, 2018 pp. 41-51

KELTNER, D. and HAIDT, J., 1999. Social Functions of Emotions at Four Levels of Analysis. *Cognition and Emotion*, 13(5), pp. 505-521.

KOOPMAN, C., 2018. The Will, the Will to Believe, and William James: An Ethics of Freedom as Self-Transformation. *Journal of the History of Philosophy*, 55(3), pp. 491-512.

KREITCHMANN, R.S., ABAD, F.J., PONSODA, A., NIETO, M.D., and MORILLO, D., 2019 Controlling for Response Biases in Self-Report Scales: Forced Choice vs. Psychometric Modelling of Likert Items. *Frontiers in Psychology* October 2019 Vol.10 Article 2309

KUHN, E., MULLER, S., HEIDBRINK L., and BUYX A., 2020 The Ethics of Workplace Health Promotion *Public Health Ethics* Vol.13 Issue 3 2020 pp. 234-246

KUHN, T.S., 1962 *The Structure of Scientific Revolutions: Second Edition* 1970. *International Encyclopedia of Unified Science.* Eds. Otto Neurath, Rudolph Carnap, Charles Morris.

LANDER, E.S., 2016 Heroes of CRISPR *Cell* 164, January 14, 2016, pp. 18-28

LE PAGE, M., 2018, 28 November 2018. CRISPR babies: more details on the experiment that shocked the world. *New Scientist Health*.

LEDFORD, H., 2016. CRISPR concerns: UK bioethics panel eyes the implications of gene editing. *Nature*, 538, pp. 17.

LEONG, M., OLNICK, S., AKMAL, T., COPENHAVER, A., and RAZZAK, R., 2016 How Islam Influences End-of-Life Care: Education for Palliative Care Clinicians. *Journal of Pain and Symptom Management* Vol.52 No.6 December 2016 pp. 771-774

LOEWENSTEIN, G., BRYCE, C., HAGMANN, D. and RAJPAL, S.D., 2015. Warning: you are about to be nudged. *The behavioral science & policy association*, (Spring), pp. 35-42.

LOVALLO, D. and KAHNEMAN, D., 2003. Delusions of success: how optimism undermines executives' decisions. *Harvard Business Review*, July issue. Available on 30 June. *Harvard Business Review*, (July issue, available on June 30).

MACKAY, N., 2020 *The Herald Online* 2nd June 2021 Neil Mackay's Big Read: The dawn of transhumanism - conversations with Dr David Eagleman, the scientist creating a world of real superhumans in his laboratory. Available at: <https://www.heraldsotland.com/news/19401897.neil-mackays-big-read-dawn-transhumanism---conversations-dr-david-eagleman-scientist-creating-world-real-superhumans-laboratory/> [Accessed September 2021]

MANDIP, K.C., and CLIFFORD J, S., 2019 A new era of gene editing for the treatment of human diseases. *Swiss Medical Weekly* 27th January 2019 149w20021 pp. 1-13

MARTEAU, T., OGILVIE, D., ROLAND, M., SUHRCKE, M. and KELLY, M., 2011. Judging Nudging: Can Nudging Improve Population Health? *British Medical Journal*, 342, pp. 263-275.

MCSTAY, A., 2020 Emotional AI, soft biometrics and the surveillance of emotional life: An unusual consensus on privacy. *Big Data and Society* January-June 2020: pp. 1-12

MELE, A., 2002 *Autonomy, Self-Control, and Weakness of Will in Kane*, 529–548.

MENARD, J. 2010 'Nudge' for Public Health Ethics: Libertarian Paternalism as a Framework for Ethical Analysis of Public Health Interventions? *Public Health Ethics* Vol. 3 No. 3 pp. 229–238

METZ, T. 2013 Ch3 *The Bearer of Meaning in 'Meaning in Life'*. Published to Oxford Scholarship Online January 2014

MILES, A., MCLOUGHLIN, M., and POLYCHRONIS, A., *Medicine and evidence: knowledge and action in clinical practice*. *Journal of Evaluation in Clinical Practice* 13 (2007) pp. 481–503

MILL, J S., 1966. A selection of his works. Springer pp. 1–147. Available at: https://scholar.google.com/scholar_lookup?title=A+selection+of+his+works&author=JS+Mill&publication_year=1966& [Accessed January 2021]

MILL, J., 2001. *On Liberty (1859)* Kitchener, Ontario: Batoche Books.

MILLER, F., and GELINAS, L., 2013 Nudging, Autonomy, and Valid Consent: Context Matters. June 2013 *The American Journal of Bioethics* 13(6): pp. 12-30

MINDSPACE Report: Influencing behaviour through public policy. June 2010. United Kingdom: Cabinet Office Behavioural Insights Team & the Institute for Government.

MONGIN, P., and COZIC, M., 2018 Rethinking Nudge: not one but three concepts. *Behavioural Public Policy* (2018), 2: 1, pp. 107–124

MORE, M., 2010. The Overhuman in the Transhuman. *Journal of Evolution and Technology*, 21(1).

MURRAY, C. J. L., and LOPEZ, A., 1996 *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020: Summary*. Cambridge: World Health Organisation

NANCY, J., 2002. *L'Intrus*. Éditions Galilée, 2000. English translation edn. Michigan State University Press.

National Advisory Council on Bioethics 2015 Department of Health Ireland Report on Nudging in Public Health – An Ethical Framework by the National Advisory Council on Bioethics

NHS Blood and Transplant 2012 Annual Review 2021-13, Available at: https://nhsbt.dbe.blob.core.windows.net/umbraco-assets-corp/1881/nhsbt_annual_review_2012-13.pdf [Accessed September 2021]

NHS Blood and Transplant 2019 All about opt out. February 2019- last update Available at: <https://www.organdonation.nhs.uk/all-about-opt-out/2019>. [Accessed February 2019]

NHS Blood and Transplant 2020 Progress Report for BAME Community Investment Scheme Initiative. Available at: <https://www.nhsbt.nhs.uk/cisreport/> [Accessed September 2021]

NHS Blood and Transplant 2021 Business Plan 2021-22 Available at: <https://www.nhsbt.nhs.uk/who-we-are/performance-and-strategy/> [Accessed September 2021]

NHS Blood and Transplant 2021a Organ Donation and Transplantation 2030: Meeting the Need. Available at: <https://www.odt.nhs.uk/odt-structures-and-standards/key-strategies/meeting-the-need-2030/> [Accessed September 2021]

NICKLES, T., 2017. Scientific Revolutions. *The Stanford Encyclopedia of Philosophy*. ZALTA, EN., ed. [Online] Winter 2017 edn. Available at: <https://plato.stanford.edu/archives/win2017/entries/scientific-revolutions/> [Accessed September 2021]

NIE, J., XIE, G., CHEN, H. and CONG, Y., 2020. Conflict of Interest in Scientific Research in China: A Socio-ethical Analysis of He Jiankui's Human Genome-editing Experiment. *Journal of Bioethical Inquiry*, 17, pp. 191-201.

NIINILUOTO, I., 2017. Optimistic realism about scientific progress. *Synthese*, 194, pp. 3291-3309.

-
- NORMILE, D., 2019, December 30, 2019, 8:15 AM. Chinese scientist who produced genetically altered babies sentenced to 3 years in jail. Science Asia/Pacific Scientific Community.
- Nudging in Public Health – An Ethical Framework. A Report by the National Advisory Committee on Bioethics. December 2015. United Kingdom: Department of Health.
- Nudging in Public Health – An Ethical Framework. A Report by the National Advisory Committee on Bioethics. Department of Health [Paraphrased from Institute for Government and the Cabinet Office, MINDSPACE: Influencing Behaviour Through Public Policy, March 2010]. 2015. (December), pp. 3-4.
- Nuffield Council on Bioethics 2018 Genome Editing and Human Reproduction: social and ethical issues (London: Nuffield Council on Bioethics)
- NUSSBAUM, M. C., 2004. Mill between Aristotle & Bentham Daedalus, Spring, 2004, Vol. 133, No. 2, On Happiness (Spring, 2004), pp. 60-68
- NUSSBAUM, M. C., 2013 Political Emotions: Why love matters for justice. The Belknap Press of Harvard University Press
- NUSSBAUM, M. C., 2018 The Monarchy of Fear: A philosopher looks at our political crisis Simon & Schuster (July 30, 2019)
- ORR, R.D., and SIEGLER, M. 2002 Is posthumous semen retrieval ethically permissible? Journal of Medical Ethics 2002;28: pp. 299–303
- PARENS, E., 2013 On Good and Bad Forms of Medicalization. Bioethics Vol.27 No.1 2013 pp. 28–35
- PASCAL, B., 1995. Pensées (1670) New York: Penguin.
- PAUL, L., 2014 Transformative Experience. Oxford University Press.
- PEDWELL, C., 2018 Habit and the Politics of Social Change: A Comparison of Nudge Theory and Pragmatist Philosophy Body and Society Vol. 23(4) pp. 59–94
- PETTICREW, M., MAANI, N., PETTIGREW, L., RUTTER, H., and VAN SCHALKWYK, M.C., 2020 Dark Nudges and Sludge in Big Alcohol: Behavioral Economics, Cognitive Biases, and Alcohol Industry Corporate Social Responsibility. The Milbank Quarterly, Vol. 00, No. 0, 2020 pp. 1-39
- PILSCH, A., 2017 Transhumanism: Evolutionary Futurism and the Human Technologies of Utopia University of Minnesota Press
- PRAINSACK, B., 2017 Personalised Medicine: Empowered Patients in the 21st Century? May 2017 New York: New York University Press.
- PUTNAM, H. 2012 Ch1 Brains in a Vat in 'Reason, Truth and History, Cambridge University Press 1981, Online Publication June 2021. pp. 1-21
- PYSZCZYNSKI, T., 2004. What are we so afraid of? A terror management theory on the Politics of Fear. Social Research, 71(4), pp. 827.

-
- RADOILSKA, L. V., 2009 Public Health Ethics and Liberalism. *Public Health Ethics*. Vol. 2 No. 2 2009 pp. 135–145
- RADOILSKA, L. V., 2010. An Aristotelian Approach to Cognitive Enhancement. *Journal of Value Inquiry*, 44, pp. 365-375.
- RADOILSKA, L.V., 2012 Akrasia and Ordinary Weakness of Will. *Tópicos*, 43. pp.25-50
Available at:
<http://www.radoilska.com/uploads/2/5/9/0/25906976/top43_radoilska_akrasia.pdf>
[Accessed June 2020]
- RANDHAWA, G., 2011 Achieving equality in organ donation and transplantation in the UK: challenges and solutions. A Race Equality Foundation Briefing Paper June 2011 Available at: <https://raceequalityfoundation.org.uk/wp-content/uploads/2018/03/health23-3.pdf> [Accessed September 2021]
- RATCLIFFE, M., 2009. The Phenomenology of Mood and the Meaning of Life. In: P. GOLDIE, ed, *The Oxford Handbook of Philosophy of Emotion*. Oxford University Press.
- RAWLS, J 1999 *Justice as Fairness in Collected Papers*, S. Freeman (ed.), Cambridge, MA: Harvard University Press, 1999.
- REARDON, S., 2019 Pig brains kept alive outside bodies for hours after death. *Nature Online* 2019. Available at: <https://www.nature.com/articles/d41586-019-01216-4> [Accessed September 2021]
- REGALADO, A., 2019 China's CRISPR babies: Read exclusive excerpts from the unseen original research. *MIT Technology Review Biotechnology/CRISPR*. December 3rd 2019
- REISCH, G.A., 1991 Did Kuhn Kill Logical Empiricism? *Philosophy of Science* Vol.58, No.2 (Jun., 1991), pp. 264-277
- ROBERGS, R.A., 2017 Lessons from Popper for science, paradigm shifts, scientific revolutions and exercise physiology. *British Medical Journal Open Sport & Exercise Medicine* 2017:3 pp. 1-11
- ROESER, S., 2010 Intuitions, emotions and gut reactions in decisions about risks: towards a different interpretation of 'neuroethics' *Journal of Risk Research*, 13:2, pp. 175-190
- ROESER, S., 2006 The Role of Emotions in Judging the Moral Acceptability of Risks. *Safety Science* 44 pp. 689-700
- ROBINS, R. and BEER, J., 2001. Positive illusions about the self: short-term benefits and long-term costs. *Journal of Personality and Social Psychology*, 80, pp. 340-352.
- SAGHAI, Y., 2013. Salvaging the concept of a nudge. *Journal of Medical Ethics*, 39, pp. 487-493.
- SAMPLE, I., 2019, Tue 31 Dec 2019 00.23GMT. Chinese scientist who edited babies' genes jailed for three years. *The Guardian News/Asia Pacific/Gene Editing*.
- SANDERS, A., and BOSTROM, N., 2006 Converging Cognitive Enhancements. *Annals New York Academy of Sciences* 1093: pp. 201–227

SAUL, J.M. 2017 Racial Figleaves, the Shifting Boundaries of the Permissible, and the Rise of Donald Trump. *Philosophical Topics*, 45 (2). pp. 97-116.

SHIH, G. and JOHNSON, C.Y., 2018 Chinese genomics scientist defends his gene-editing research in first public appearance. *The Washington Post Asia & Pacific*. November 28th 2018

SKERRETT, P., 2015 Experts Debate: Are we playing with fire when we edit human genes? In: *STAT online* 17th November 2015, Available at: <https://www.statnews.com/2015/11/17/gene-editing-embryo-crispr/> [Accessed September 2021]

SKINNER, B., 1971. *Beyond Freedom and Dignity*. Penguin Books.

SLABY, J., 2010. Steps towards a Critical Neuroscience. *Phenomenology and the Cognitive Sciences*, 9(3), pp. 397-416.

SLOVIC, P., 1987 Perception of Risk. *Science*, VOL. 236 May 1987 pp. 280-285

SOLBERG, C.T., SORHEIM, P., MULLER, K.E., GAMLUND, E., NORHEIM, O.F., BARRA, M., 2020 The Devils in the DALY: Prevailing Evaluative Assumptions. *Public Health Ethics*. Vol.13 Issue 3 pp. 259-274

SOLOMON, M., 2015. *Making Medical Knowledge*. Oxford University Press.

SPARROW, R., 2019 Yesterday's Child: How Gene Editing for Enhancement Will Produce Obsolescence—and Why It Matters. *The American Journal of Bioethics*, 19:7, pp. 6-15

SPROKKEREEF, A., DE HERT, P., 2007 Ethical Practice in the Use of Biometric Identifiers Within the EU. *Law, Science and Policy*, 2007, Vol. 3, pp. 177–201

STANLEY, J., 2015 *How Propaganda Works*. Princeton. Princeton University Press.

STROUD, S. and TAPPOLET, C. (eds.), 2003, *Weakness of Will and Practical Irrationality*, Oxford: Clarendon Press

SUNSTEIN, C., 2001. Probability Neglect: Emotions, Worst Cases, and Law. *Chicago Law School, Coase-Sandor Working Paper Series in Law and Economics (Working Paper No.138)*.

SUNSTEIN, C. 2019 *How Change Happens*. The MIT Press, Cambridge MA. London

TAPPOLET, C., 2009. Emotion, Motivation, and Action: The Case of Fear. In: P. GOLDIE, ed, *The Oxford Handbook of Philosophy of Emotion*. Oxford University Press.

THALER, R. and SUNSTEIN, C., 2008. *Nudge: Improving Decisions about Health, Wealth and Happiness*. New Haven, CT: Yale University Press.

THALER, R. and SUNSTEIN, C., 2010. *The Nudge* [blog]. <http://nudges.org/2010/06/02/assorted-links-33/>; J. BALZ, ed, [June 2010]

The European Commission's compiling agency: Eurostat, the statistical office of the European Union Glossary, 2021 Available at: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Standardised_death_rate_\(SDR\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Standardised_death_rate_(SDR)) [Accessed September 2021]

TIMMINS, N., 2015. Discussion- Optimism about the future. *Health Economics, Policy and Law*, 10, pp. 107-111.

TUERLINGS, E., 2019 WHO Expert Advisory Committee on Developing Global Standards for Governance and Oversight of Human Genome Editing. Background Paper Governance 1 Human Genome Editing. Commissioned by the World Health Organization to serve as a background document for the first meeting of the Expert Advisory Committee on Developing Global Standards for Governance and Oversight of Human Genome Editing (18-19 March 2019).

TVERSKY, A., and KAHNEMAN, D., 1981 The Framing of Decisions and the Psychology of Choice. *Science* Vol.211 pp. 453-458

UFFELMAN, M., 2011. Forging the Self in the Stream of Experience: Classical Currents of Self-cultivation in James and Dewey. *Transactions of the Charles S. Peirce Society*, 47(3), pp. 319-339.

VAN DEURSEN, J., 2014. The role of senescent cells in ageing. *Nature*, 509 (22 May), pp. 439-446.

WARTELLA, E., LICHTENSTEIN, A., YAKTINE, A. and NATHAN, R., 2011. *Front-of-Package Nutrition Rating Systems and Symbols: Promoting Healthier Choices*. The National Academies Press.

WATSON, G., 1977, "Skepticism about Weakness of Will," *Philosophical Review*, 85: 316–39

WEIJERS, D., 2014. Optimistic Naturalism: Scientific Advancement and the Meaning of Life. *Sophia*, 53, pp. 1-18.

Wellcome Global Monitor – First Wave Findings. How does the world feel about science and health? 2019. Gallup.

WELLS, D., 2008 Genetic Engineering in Athletes *British Medical Journal* 2008 Jul 12; 337(7661): pp. 63–64

What is Independent SAGE 2021 Available at:

<https://www.independentsage.org/independent-sage/> [Accessed June 2021]

WILLIAMS, B., 1973. *The Makropoulos Case: reflections on the tedium of immortality*. Problems of the Self. Cambridge University Press, ch6.

WOLF, S., 2007. *Meaningfulness- A third dimension of the good life*. Tanner Lectures, Princeton, (Condensed version, Brown University).

World Health Organisation 2021 4th April Covid-19 Strategy Update Available at:

<https://www.who.int/publications/m/item/covid-19-strategy-update> [Accessed September 2021]

ZEALLEY, B. and DE GREY, A., 2012. Strategies for Engineered Negligible Senescence. *Gerontology*, (October).

ZIMMER, K., 2018 CRISPR Scientists Slam Methods Used on Gene-Edited Babies. *The Scientist News & Opinion*. December 4 2018

Appendix

Fig 1.1 Donors after brain death UK numbers 2018-2021

	2020-2021*	2019-2020	2018-2019
Donors after brain death (DBD)	691	946	961
Donors after circulatory death (DCD)	360	638	639
Total deceased donors (DD)	1051	1584	1600
Non-proceeding deceased donors (DD)	406	719	678
Living donors (LD)**	355	970	958

* 2018-2019 and 2019-2020 data run from 31st March to 1st April. 2020-2021 data was available for the period 31st March 2020 to January 2021.

These numbers were taken from documents made available by NHS Blood and Transplant, on their public website. 2018-2019 and 2019-2020 figures from 'Annual activity figures, April 2020', 2020-2021 figures from 'Monthly Statistics, Feb 2021'.

**LD data are recorded one month in arrears.

Fig 1.2 'If you believe in donation, prove it'

NHS Blood and Transplant



Fig 1.3

'Be a Hero' (NHS Yorkshire)



Fig 1.4

'Kill Jill' (2009) Life and Death campaign from © 2019 The Union Advertising Agency Ltd for the Scottish Government

Healthier Scotland



Fig 1.5

'If you needed an organ transplant, would you have one?'

Yes I donate, Organ donation

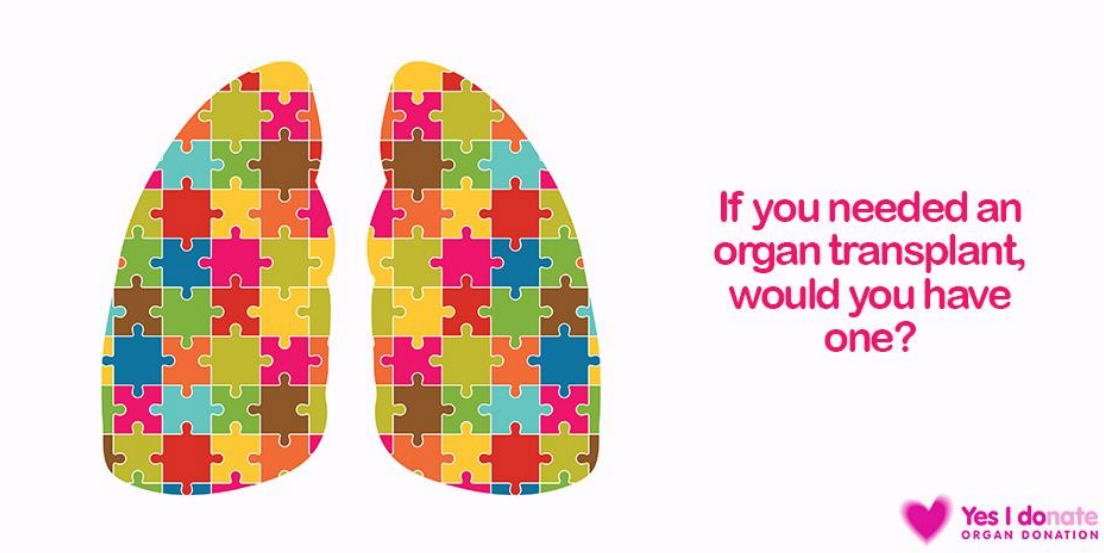


Fig 1.6

'O Negative can be given to anyone'

Save a life, give blood



Fig 1.7

'Become a blood donor'

Save a life, give blood

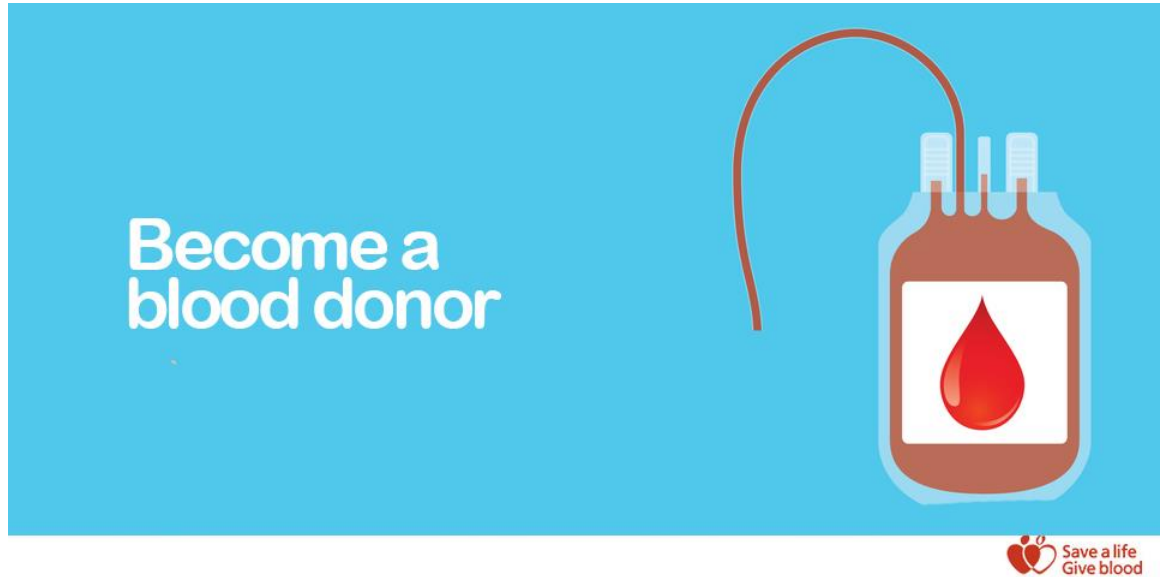


Fig 1.8

MINDSPACE

Behavioural Insights Team report, MINDSPACE Influencing behaviour through public policy, published by the Cabinet Office 2010 pp. 8

Messenger	we are heavily influenced by who communicates information
Incentives	our responses to incentives are shaped by predictable mental shortcuts such as strongly avoiding losses
Norms	we are strongly influenced by what others do
Defaults	we "go with the flow" of pre-set options
Salience	our attention is drawn to what is novel and seems relevant to us
Priming	our acts are often influenced by sub-conscious cues
Affect	our emotional associations can powerfully shape our actions
Commitments	we seek to be consistent with our public promises, and reciprocate acts
Ego	we act in ways that make us feel better about ourselves