

# An evaluative conservative case for biomedical enhancement

John Danaher

## Correspondence to

Dr John Danaher, School of Law, NUI Galway, University Road, Galway, Ireland; johndanaher1984@gmail.com

Received 11 December 2015

Revised 12 May 2016

Accepted 7 June 2016

Published Online First

27 June 2016

## ABSTRACT

It is widely believed that a conservative moral outlook is opposed to biomedical forms of human enhancement. In this paper, I argue that this widespread belief is incorrect. Using Cohen's evaluative conservatism as my starting point, I argue that there are strong conservative reasons to prioritise the development of biomedical enhancements. In particular, I suggest that biomedical enhancement may be essential if we are to maintain our current evaluative equilibrium (ie, the set of values that undergird and permeate our current political, economic and personal lives) against the threats to that equilibrium posed by external, non-biomedical forms of enhancement. I defend this view against modest conservatives who insist that biomedical enhancements pose a greater risk to our current evaluative equilibrium, and against those who see no principled distinction between the forms of human enhancement.

## INTRODUCTION

Conservatives like things the way they are; enhancement technologies allow us to change them. It is no surprise then that there is a widespread belief that conservative moral principles lead to the rejection of human enhancement. In this article, I argue that this widespread belief is wrong in a number of critical respects. I argue that the promotion of biomedical human enhancement could be crucial if we are to conserve our current evaluative equilibrium against an array of technological threats. I defend this argument in four parts. First, I situate my argument within both the conservative moral tradition and the enhancement debate. Second, I present and clarify the argument itself. Third, I defend the central premise of the argument, noting three specific ways in which biomedical enhancement could help to preserve our current evaluative equilibrium. And finally, I respond to six objections to this argument.

## ENHANCEMENT AND EVALUATIVE CONSERVATISM

It is important to clarify the style of conservatism adopted throughout the remainder of the article and to explain how the argument I defend is situated relative to other, somewhat similar, arguments.

Conservatism is a rich philosophical tradition. My argument draws on a thin, but core, version of this tradition. The version has two components: an evaluative one and an epistemic one. The evaluative component maintains that we should seek to hold onto or preserve existing sources of value *even if we are promised newer better sources of value in their stead*. This evaluative conservatism is taken from the work of Jerry Cohen.<sup>1 2</sup> Cohen's

conservatism holds that there is reason to favour an existing source of value ( $V_1$ ) over a potential/future value ( $V_2$ ), even if there is also reason to suspect that the (expected) value of  $V_2$  would be greater than  $V_1$ . An example helps.<sup>3</sup> Suppose you currently have a pet dog—call him Fido. You have established a deep and affectionate relationship with Fido. One day someone offers you a new dog—call him Rex. You are assured (and you believe) that you could have an even deeper and more affectionate relationship with Rex, but only if you give up Fido. Presented with such a choice, many people would hang onto Fido. Why? Because they are evaluative conservatives: They wish to continue to value that which is valuable and valued by them,<sup>i</sup> partly because of the thing is a source of value and partly because of the personal and/or cultural history they associate with that source of value. In other words, they want to hold onto existing sources of value because of how those sources figure in their personal and cultural identity, in preference to alternative and arguably better sources of value. My conservative case for biomedical enhancement works from this evaluative form of conservatism. It maintains that we should favour biomedical enhancement because doing so will help us to preserve existing personal, social and political sources of values that are currently under threat. In doing so, it also appeals to our personal and cultural attachment to these sources of value.<sup>ii</sup>

In addition to the evaluative component, I include an epistemic component in the conservative stance adopted in this article. This helps to situate my argument within the popular Burkean tradition of conservatism.<sup>5</sup> This form of conservatism is sceptical of change on the grounds that existing social arrangements can embody a degree of distributed knowledge that is difficult to make accessible to conscious awareness and deliberation, and hence there is significant evaluative risk in seeking to change from these existing arrangements. In this sense, the argument presented could be linked with the concerns expressed

<sup>i</sup>Pugh *et al*<sup>3</sup> discuss the distinction between valuing the valuable and valuing the valued. The former is about intrinsic value, the latter is about personal history with the source of value.

<sup>ii</sup>A reviewer pointed out that evaluative conservatism could have paradoxical consequences. What if, for example, one currently values constant change? If so, then one would constantly favour changes to the existing status quo, which hardly seems conservative. Interestingly, this is a criticism of enhancement that some modest conservatives develop.<sup>4</sup> This might be logically permissible, but it seems unlikely that anyone espousing a conservative moral outlook would favour this view, hence, I ignore it in what follows.



CrossMark

To cite: Danaher J. *J Med Ethics* 2016;**42**:611–618.

in debates about the role of moral risk in practical reason, though I will not spell out those links in what follows.<sup>6–9</sup> The epistemic component is relevant because although Cohen's evaluative conservatism works even in a situation in which we know for sure that the future source of value will be better than what we currently have, it seems more realistic to assume that we make our moral choices in the face of risk and uncertainty.

I make no further conservative commitments in this article. I would suggest, though I will not defend this suggestion, that the evaluative and epistemic components that I have identified represent the core of most forms of political and moral conservatism. But to this core additional commitments are often added. For instance, bioconservatives often make strong assumptions about the fixity and imperfections of human nature/biology and the desirability of a natural order.<sup>10</sup> I deliberately eschew those commitments here. I do so for two reasons. The first is that these thicker and more traditional forms of conservatism have been widely defended and critiqued already.<sup>10–12</sup> The thinner and more modest forms of conservatism are more novel and underexplored. Second, the thinner form of conservatism is arguably more formidable: by being modest in its conservative commitments it tries to show how conservative values can appeal across the political and moral spectrum. Cohen, for instance, was no political conservative but he found evaluative conservatism appealing.

Proponents of this thinner form of evaluative conservatism have used it to argue against human enhancement technologies. Cohen believed, and others have agreed, that evaluative conservatism provided powerful reasons to oppose enhancement.<sup>3</sup> Likewise, recent anti-enhancement authors such as Nicholas Agar and Michael Hauskeller,<sup>4 13–15</sup> although not explicitly calling themselves conservatives, have challenged the enhancement project on the grounds that it threatens existing sources of value. Agar worries that enhancement technologies prioritise hypothetical external goods over existing internal (and species-specific) goods. Hauskeller explicitly laments the fact that enhancement favours the possible better over the existing good.<sup>4</sup> My goal in this article is to suggest that evaluative conservatism actually provides powerful reasons to favour human enhancement. It does so because enhancement can help us to preserve much of what is currently valuable about our way of life.

This basic insight is not novel. Pugh *et al*<sup>3</sup> have pointed out that Cohen's conservatism can be used to defend some forms of enhancement. For instance, they argue that life extension therapy can be defended on the grounds that it preserves the value that attaches to our current lives and, similarly, mood enhancement drugs can be defended on the grounds that they allow us to preserve the value of existing interpersonal relationships. Buchanan makes similar claims, pointing out that “[t]o avoid losing some of the good things we now enjoy, we will have to enhance ourselves in particular ways” (ref. 12, p. 163). But none of these authors develops the conservative case beyond a mere suggestion. They shine a light down a possible pathway; they do not lead us to the end. That is what I want to do in what follows. I want to argue that evaluative conservatism can lead to the view that biomedical enhancement is close to an axiological imperative: something we must pursue lest we lose much of what is currently valuable (and valued by us) in our political, economic and personal lives.

### THE EVALUATIVE CONSERVATIVE ARGUMENT

The argument is that *biomedical enhancement* may be essential if we are to conserve our current *evaluative equilibrium*. This argument appeals to a number of controversial concepts, each of which needs to be clarified.

I start with the most obvious point of contention: the idea of *enhancement* itself. There are two prominent approaches to the definition of enhancement in the literature.<sup>16–20</sup> There is the *functionalist* approach, which holds that an intervention counts as an enhancement if it improves (or adds to) the functionality of a human being relative to some population-level (or species-level) norm.<sup>10</sup> And there is the *welfarist* approach which holds that an intervention counts as an enhancement if it increases the chances of someone living a good life, relative to their current circumstances.<sup>16</sup> The argument I defend works best with the functionalist definition, though it will also work with the welfarist definition, provided that the ‘good life’ is understood to include more than simply subjective well-being and, in particular, to include some sense of personal achievement, mastery and understanding.

The second point of contention is the focus on *biomedical* forms of enhancement. As Buchanan<sup>10</sup> notes, enhancement is a broad concept and could refer to interventions as diverse as farming, democratic governance, literacy, education, pharmacological drugs, brain stimulation and neuroprosthetic devices. Some people insist that there are no principled distinctions between these forms of enhancement.<sup>21</sup> Others insist that there are.<sup>22</sup> In this article, I use the descriptor ‘biomedical’ to designate enhancement technologies that either directly target (eg, pharmacological drugs) or directly integrate with (eg, neuroprosthetics) our human biology. They are to be contrasted with external and automating technologies, which often make our lives easier, may indirectly increase our power to change the world around us, and so could also be said to have an enhancing effect. I will insist that there is an important distinction between the two types of technological intervention when it comes to the conservation of value.

This brings me to the third and final point of contention, which relates to my claim that biomedical enhancement can conserve our current *evaluative equilibrium*. I appeal to this concept in order to incorporate Cohen's evaluative conservative principles into my argumentative framework. I define ‘evaluative equilibrium’, loosely, to mean the set of values that undergirds and permeates our current way of life. In what follows, I particularly appeal to the set of values that undergird the dominant Western political and economic systems, and those that support our sense of personal achievement and satisfaction. Evaluative conservatism then enters the picture as an attempt to defend that current evaluative equilibrium in lieu of alternative evaluative equilibria promised or predicted by certain technological enthusiasts (even if those evaluative equilibria promise to be better). I appreciate that this notion of a readily identifiable evaluative equilibrium is questionable, but I hope to add some reasonably specific content to it below.

This brings us to the argument itself, which can be formulated like this:

1. We ought to maintain our current set of values ( $VS_1$ ) over a potential/future set of values ( $VS_2...VS_n$ ), even if those future sets of values are potentially superior.
2. Prioritising biomedical enhancement can help us to maintain  $VS_1$  and stave off threats to  $VS_1$  posed by other technological developments.
3. Therefore, we ought to prioritise biomedical enhancement.

There is an unavoidable degree of empirical uncertainty in premise (2).<sup>iii</sup> This is why the argument talks about how

<sup>iii</sup>On the role of empirical uncertainty in the enhancement debate, see Loi.<sup>23</sup>

biomedical enhancement ‘can help’ rather than ‘will help’, and why the conclusion is framed in terms of ‘prioritisation’ and not actual use. Nevertheless, the argument defends a reasonably strong view about the value of biomedical enhancement. It claims that the development of biomedical enhancement technologies could be a moral imperative due to the threat to our current evaluative equilibrium that is posed by other technological developments. Hence, it is not simply that biomedical enhancement is desirable or preferable. The strength of this should become clearer in the next section.

For the purposes of this argument, I assume that the conservative principle stated in premise (1) is correct. Readers will need to go elsewhere for its defence.<sup>1–3</sup> My focus thus switches to the defence of premise (2). I defend that over the remaining two sections of the article.

### THE THREAT OF EXTERNALISATION TO OUR EVALUATIVE EQUILIBRIUM

Proponents and opponents of enhancement often consider the ethical implications of enhancement technologies in isolation from other technological and social trends.<sup>iv</sup> Thus, for example, they might consider the impact of cognitive enhancement drugs among university students against the more or less static socio-normative background of contemporary (often Anglo-American) universities. This allows them to identify ways in which the use of such drugs may be desirable or undesirable, given that background. But it does not enable them to consider how the use of such drugs may fit into an evolving and ever-changing model of university life. The central theme of my argument is that this sort of static ethical analysis is a mistake. When considering the ethical implications of biomedical enhancement, we must take into account the dynamic and ever-changing sociotechnical background in which such enhancements are likely to develop. It is when we do this that we see most clearly the conservative value of such technologies.

This means that the first thing we need to do in order to defend premise (2) is to consider the background of other technological changes against which the assessment of biomedical enhancement must take place. It is impossible to provide a complete description of that background, but we can at least highlight some of its most salient features. I would argue that the biggest and most important feature of that background is the rapid growth in computerisation, artificial intelligence, machine learning, robotics and other types of information technology. I do not view this as a particularly controversial claim. The growth in these technologies is readily transparent. It has been exponential for much of the past 60 years and some are convinced that this exponential growth will continue, and will have spillover effects as information communication technology ‘infects’ other technological developments.<sup>25–27</sup> There are three important manifestations of this trend: (i) the rise of automating technologies in agricultural, manufacturing, professional and service-related work;<sup>28–31</sup> (ii) the increasing use of data-mining and other automated decision-making technologies in governance systems<sup>32–33</sup> and (iii) the use of digital assistants and robot helpers in our personal lives.<sup>33–34</sup> My claim is that these three

technological developments pose a threat to our current evaluative equilibrium. Once we see this, we can see the conservative potential of biomedical enhancements. It is best to do this by addressing the threats individually.

I start with the economic threat. Most Western societies are characterised by mixed economic systems in which the government provides legal and administrative support for a system of private capitalistic enterprise. Most people are employed (or hope to be employed) within that system. These employments are a significant source of value to the individuals being employed and to society more generally. It is how goods and services get distributed to those who need or desire them. It is how individuals secure the economic rewards that make life possible and pleasurable. Many people tie their self-worth and self-actualisation to their success in their working lives, and are then socially evaluated on the basis of this success. Welfare systems provide some sort of safety-net, but this is usually insufficient and typically tied to the ability to work. Furthermore, welfare systems do not compensate for the non-monetary goods of work. As Ghaeus and Herzog have recently argued, work is a privileged context for the attainment of goods such as excellence/mastery, social contribution, sense of community and social status.<sup>35</sup> As such, working for economic reward forms a significant part of our current evaluative equilibrium. The problem with the rise of automation technologies is that they threaten to undermine this system of values. An increasing number of authors are convinced that these technologies will displace much of the current workforce, thereby depriving them of this source of value.<sup>28–29</sup> Furthermore, even those who are not convinced by the strong technological unemployment thesis seem to agree that automating technologies change the types of work that are available, making low-paid, precarious, but difficult-to-automate work more common.<sup>36</sup> Others even go so far as to draw the connection between this trend and the enhancement debate and argue that technological unemployment is likely to have a ‘disenhancing’ effect, at least in the welfarist sense of the term.<sup>23</sup>

Some welcome the disrupting effect of automation on work, but they are typically radicals who long for a ‘postwork’ future in which we have abandoned the ideological glorification of the work ethic.<sup>37</sup> There is little solace in this radical posture for the evaluative conservative who sees the current system as a source of value, one to which they have grown attached. The person who has worked all their lives in a profession like law and developed mastery over a certain set of valuable skills (eg, legal research, document review, brief writing) is going to lose that source of value if all those skills can be automated. Even if the promised postwork future would be better for this person, the evaluative conservative would insist that the person is axiologically warranted in trying to hold onto what they currently have. Others suggest that a basic income guarantee could address some of the problems posed by technological unemployment.<sup>28–29–38</sup> It would enable everyone to receive an income irrespective of their ability or willingness to work. This means some of the consumption-related value of work could be conserved. But the problem with this is that the non-monetary values that attach to work would not. So a basic income may partly address the concerns of the evaluative conservative, but not all of them. Biomedical enhancement, on the other hand, is more hopeful. Biomedical enhancements could allow human beings to ‘keep pace’ with their automated rivals, either through the direct augmentation of biological systems, or through the

<sup>iv</sup>Some discussions do not omit this point. A good example is the work of Persson and Savulescu<sup>24</sup> on moral enhancement, which is explicitly framed in terms of broader climatological and technological changes.

grafting in of prosthetic devices. Biomedically enhanced humans could continue to be valued contributors to the economic system. This could help to conserve work as a source of value.<sup>v</sup>

Consider a simple example. A surgeon has dedicated their lives to mastering a set of highly precise motor skills and marrying that to a deep understanding of how the human body works. Their work is individually meaningful and socially valuable. Unfortunately, the hospital has decided to invest in robotic surgeons that are more precise, more efficient, less prone to fatigue and automatically updated with the latest medical knowledge. The surgeon is told that their work is no longer needed. The result might be improved medical outcomes for the hospital, but a significant loss in meaning and value for the surgeon whose identity is so closely tied to their competence at work. This outcome could have been avoided if the hospital (or whichever relevant authority) prioritised biomedical enhancement of human surgeons over investment in robotic surgeons. Human surgeons could be augmented with enhancing drugs that make their hands steadier, improve their concentration, reduce their fatigue and enable them to better absorb the latest medical knowledge. At least in principle, there was a way to achieve better outcomes while still conserving the value of work to the surgeon. There were choices to be made about investment and prioritisation of technological innovation that could have retained our existing work-related evaluative equilibrium.<sup>vi</sup>

The second threat arises in the political sphere. Most Western societies are characterised by liberal democratic systems of governance. These systems of governance are undergirded by the commitment to moral equality and individual autonomy.<sup>39</sup> These commitments imply that governance structures are illegitimate unless they satisfy certain conditions.<sup>39 40</sup> These legitimacy conditions come in two major flavours: instrumentalist and proceduralist. Instrumentalist legitimacy conditions justify governance structures in terms of their outputs, that is, do they maximise some value beloved by the polity (less crime, more healthcare, more economic growth, longer lives, better education, etc)? Proceduralist legitimacy conditions justify governance structures in terms of the processes they use to produce their outputs, that is, are these processes inclusive, participative, comprehensible, publicly accessible and so forth? Although some favour purely instrumentalist or proceduralist approaches to legitimisation, most accept that a mix of instrumentalist and proceduralist conditions must be satisfied.

But here is the rub: algorithmic and automated decision-making is increasingly being used by political governance systems.<sup>41–43</sup> These systems limit or obviate the need for human input and oversight of decision-making. Algorithms can collect and organise decision-relevant information, draw inferences from that information and use it to make important political, legal and bureaucratic decisions. Dormehl<sup>33</sup> gives the example of a facial recognition algorithm used by the Massachusetts

Registry of Motor Vehicles to detect probable false identities in its database of driver ID photographs (157–8). Once the system found a suspiciously similar match, an automatic letter was issued cancelling the driver's licence. Similar systems are being used to detect potential terrorist threats and likely tax cheats. And the number of such systems is rapidly growing.<sup>44</sup>

Their introduction is often defended on the grounds that they are more accurate, less costly and more efficient than traditional bureaucratic systems.<sup>33 45</sup> But are they consequently more legitimate? This is highly debatable. They may be more instrumentally legitimate insofar as they can facilitate better outcomes, but their procedural virtues are dubious. These systems typically rely on mass surveillance, large-scale data-mining operations and machine-learning technologies. In addition to the privacy-related concerns that these systems raise, they also raise considerable transparency, comprehensibility and accessibility concerns. The highly technical nature of the systems makes them comprehensible to a relatively limited set of human elites. And more advanced versions of these systems are often impenetrable to even these elites, due to both the opacity of the technology, and the fact that they are woven into complex ecosystems of pre-existing algorithms.<sup>46 47</sup> The result is that the increased use of algorithmic and automated systems of governance poses a threat to the procedural legitimacy of our political system, which again constitutes a central part of our evaluative equilibrium.

The threat can be explained by way of analogy. A long-standing concern in political theory is with the attractiveness of non-democratic governance. Some (eg, Plato and JS Mill) have argued that human epistemic elites (ie, humans with greater cognitive ability, education, technical knowledge and so forth) should be given more decision-making power. Others have challenged this on the grounds that deferring to human epistemic elites often means blocking ordinary people from participating in the decisions that affect their lives and/or failing to explain and justify those decisions to them. The result is something Estlund calls the threat of 'epistocracy'—a threat that undermines participative democratic governance.<sup>48</sup> The growth of algorithmic governance gives rise to a similar threat: humans create algorithmic systems in the belief that they can gather and make sense of more information than humans ever could. They start deferring to the algorithms for epistemic reasons. This may be unproblematic at first, but as the algorithmic systems become more complex, comprehensive and opaque, the humans may end up deferring to them in blind faith. The result is something like a threat of 'algocracy'—humans are blocked from participating in and understanding the algorithmic decisions that affect their lives.<sup>43</sup>

The conservative impulse might be to address this problem by limiting, banning and regulating the use of such systems. This may have some utility, but one should not underestimate the difficulties. There are already numerous 'secrecy' laws that facilitate the construction and opacity of these systems, and these would have to be dismantled at national and supranational levels. Furthermore, the technologies that allow for this are already woven into our everyday lives and so it makes sense for similarly structured governance systems to be built on top of them. There are also powerful ideological forces behind their creation and use.<sup>49</sup> Resistance and regulation may be futile in this domain. Could biomedical enhancement help? There are reasons to think that it could. First, biomedical enhancements could help to improve the capacities and efficiencies of human workers within governance systems, thereby countermanding some of the drive towards their replacement. Imagine that

<sup>v</sup>Loi (ref. 23, p. 209) makes a similar observation in his discussion of technological unemployment, arguing that biomedical enhancement could be used to counterbalance the disenchanting effect of workplace automation.

<sup>vi</sup>There are also important political and social choices to be made regarding who benefits most from the productivity gains of either enhanced or automated labour. One major concern about automated labour is that it increases the share of wealth that goes to owners of capital and thereby exacerbates income and social inequality. While biomedical enhancement of workers is unlikely to solve inequality, it may be considerably *less worse* than automation since workers can at least retain their income from work.



instead of relying on an automated facial recognition algorithm, the registry of motor vehicles encouraged its employees to adopt neural prosthetics that could augment their visual acuity and speed, and enable them to see the suspiciously similar matches. Second, biomedical enhancements that are more widely distributed could help to improve the capacities of all people affected by those governance systems. For instance, widespread cognitive enhancement, in conjunction with educational programmes, could enable people to better understand and participate in an increasingly automated governance system. The pharmacological enhancements could focus on core capacities and competencies—concentration, memory retrieval, basic reasoning capacities—and the education programme could focus on explaining the mechanics of automated governance. This would help to combat some of the proceduralist concerns raised in the previous paragraphs. These biomedical enhancements could work in tandem with regulatory interventions that give people the right to challenge and access these systems, thereby doing a better job conserving this part of the current evaluative equilibrium. So the claim here is not that biomedical enhancements solve the problem on their own, but rather that they could be an essential piece of the puzzle: without enhanced reasoning capacities people will be overwhelmed and poorly disposed towards acquiring the requisite knowledge that will allow them to conserve the value of participative democratic governance.

The final threat arises in the personal sphere, that is, the day-to-day activities that make up our lives. There is a general desire to live a ‘good life’, where this can be understood as a life of meaning and virtue. Work plays a part in this conception of the good life, but obviously work itself is not the be-all and end-all of the well-lived life. We also seek meaning and virtue in our family lives, leisure pursuits, unpaid work, charitable endeavours, sporting activities and so on. In these pursuits, a sense of personal achievement is often seen as central to whether they provide meaning or not.<sup>50–51</sup> The longing for achievement is part of the current evaluative equilibrium: achievement is a source of value and something valued by us in and of itself. But achievement is only really made possible when there is an appropriate causal-mental link between what an individual does and the outcome they are trying to achieve. The problem is that the rise of AI and robotics threatens to sever that link. This happens through the rise of personalised automation. It is already obvious that individuals rely on assistive software technologies in their everyday lives, for instance, software mapping technologies and personal digital assistants. There is also the emerging field of personal robotics, which includes robots that mow the lawn, vacuum the floors, cook your meals and so on. Reliance on such technologies can be expected to grow. This threatens to sever the link necessary for personal achievement and also gives rise to a disenchanting effect on personal cognitive capacity.<sup>52</sup> Several studies demonstrate that the use of assistive software can lead to the degeneration of the cognitive capacities needed to perform certain mental and physical tasks.<sup>53–55</sup> This can occur even if there is a residual causal-mental link between what the individual does and the outcome in the world, provided that residual link is mediated through some degenerating technology.

This threat to the evaluative equilibrium may be less severe than the previous two. Personalised automation has its merits and demerits; it all depends on how it is used. Some people could use it wisely to outsource mundane or limiting activities and spend the rest of their time on activities that really do contribute to the good life. But we should be cautious in our

optimism. People can be lazy and prone to biases that cause them to do things that are counterproductive to their own well-being. Overreliance on personalised automation could exacerbate these tendencies. This will be particularly true if the technology promises to do a better job of achieving the relevant outcome than a human ever could or if it supplies addictive hedonic rewards that distract us from those achievements.<sup>vii</sup> Biomedical enhancement can, once again, be part of the conservative solution to these problems. First, mood enhancers and cognitive enhancers could be adopted to mitigate or limit the biases that drive overreliance on personalised automation. And second, it could be used to enhance human capacities so that there is no need to turn to those technologies for better outcomes.

In summary, work, legitimate governance and personal achievement are prominent features of our current evaluative equilibrium; developments in AI, robotics and information technology threaten those features, but biomedical enhancements could be used to conserve them. There is empirical uncertainty within this argument but wariness in the face of an uncertain future is part and parcel of the conservative position favoured in this article. The argument makes a strong claim about the conservative power of biomedical enhancement. The features of our evaluative equilibrium that have been singled out by the argument are not mere sideshows. They are central to our current economic, political and social institutions. Furthermore, the threats posed by the automating technologies are not insignificant. The erosion of work, comprehensible governance or personal achievement individually constitute significant disruptions to our evaluative equilibrium. When two or more of them combine, the effect is truly radical. The evaluative conservative should be strongly committed to minimising such disruption and biomedical enhancement may, according to the arguments proffered above, be the most viable form of conservation.

## OBJECTIONS AND REPLIES

I close by considering six objections to this argument.

The first objection is: why cannot the conservative adopt a strategy of general resistance to technological development? My argument supposes that they must accept the reality of workplace automation, algorithmic outsourcing in government and increased infiltration of artificial intelligence (AI) and robotic assistance into everyday life. Surely, it is possible and preferable to resist those developments *and* resist biomedical enhancement as well.

There is some merit to this criticism. It probably would be ideal, from the conservative’s perspective, to resist all the developments since that would conserve more of the existing evaluative equilibrium. The question is whether this is feasible. The conservative may have to pick and choose their battles. When discussing the rise of algorithmic governance in the previous section, I suggested that it might be difficult to resist this phenomenon due to the fact that: (i) the underlying technologies are woven into the fabric of everyday life; (ii) there are laws and policies in place that facilitate and encourage their development

<sup>vii</sup>An extreme example would be the vision of the future depicted in the Pixar movie *Wall-E*. The imagined future consists of innumerable robots that perform routine tasks and supply humans with all their needs and pleasures. The humans are reduced to passive, overweight beings, moving about on automated chairs, receiving a constant and addictive supply of sugary foods and light entertainment. I suspect this possible future is highly unlikely to materialise, but its exaggeration of the human susceptibility to short-term rewards is instructive.

and (iii) there are powerful ideological forces driving their adoption. The same holds in the case of workplace automation. Indeed, automation in work is deeply engrained into the incentive structure of capitalism. It is less true in the personal sphere, but automation is becoming increasingly normalised there too and many of same forces underlying the drive towards workplace automation are behind the push towards personal automation. The question then becomes whether it is possible to resist the push towards automation across all three domains. It may be, but one advantage of biomedical enhancement is that it can facilitate conservation of values across all three domains at the same time. This is because success for humans across all three domains is dependent on general human intelligence and ability, and biomedical enhancement tends to target generic cognitive, emotional and perceptual capacities that are relevant to value conservation across all three domains. Until we develop generalised machine intelligence, the technologies that pose a threat to all three domains will be diverse and disparate.

The second objection is that even if biomedical enhancements preserve the values mentioned above, could not it also be the case that they change other values that the conservative favours? Sandel's giftedness argument illustrates the objection.<sup>56</sup> He argues that one prominent and important feature of our existing evaluative equilibrium is our gratitude for the gifted aspects of our lives. That is to say, gratitude for the fact that many good things happen to us because of biological and sociological factors that are beyond our control. This has a number of positive consequences in terms of empathy and social solidarity.<sup>4</sup> He worries that if we continue to pursue biomedical enhancement, we will lose this element of our evaluative equilibrium because everything will become contingent and manipulable. Similarly, there are those who worry that biomedically enhanced personal achievement is less valuable because it is less authentic.

Three things can be said in response to this. First, these forms of conservative argument are often unpersuasive on their own merits. 'Authenticity' can be parsed in different ways, sometimes as a fear about how biomedical enhancements bypass personal effort, other times as a fear about disconnection from the authentic self. But there is no reason why biomedical enhancements have to eliminate effort: augmenting a capacity does not mean you stop using it in an effortful way.<sup>57</sup> Nor is it clear why it would lead to disconnection from the authentic self: if the decision to enhance is undertaken with appropriate self-reflection then it can be an expression of the authentic self. Second, biomedical enhancements could be prioritised and selected on the basis that they help to conserve the values identified in these arguments. For instance, feelings of gratitude, empathy and solidarity are things that can be conserved and enhanced through various pharmacological and behavioural interventions.<sup>58</sup> Finally, even if this is all wrong, part of the motivation behind the argument in this article is that the evaluative conservative may sometimes face important tradeoffs when it comes to the conservation of value. The question then becomes whether the threat to something like gratitude and authenticity outweighs the three threats I have identified.

A third objection holds that there is a paradox at the heart of my argument. My claim is that externalised automation technologies pose a series of threats to our current evaluative equilibrium. This is because these technologies take over the performance of tasks traditionally performed by humans. But would things really be any better if (enhanced) humans continued to perform those tasks? It is well-known that the routine performance of a task by a human gives rise to *internal automaticity*, that is, the task is initially learned and performed through

conscious mental representations, but as the human gets better it transfers to subconscious systems within the human body. This is particularly true for many tasks performed at work, as well as certain hobbies and skills. If everything ends up automated in the end (either through external automation or internal automaticity), what difference does it make whether it is mediated through the human body?

Two responses seem apposite. First, this objection has a limited scope. While it is true that many human tasks end up being internally automatised, it is not always true. Certain cognitive tasks are always performed through conscious mental representations and their continued conscious representation may be part of what makes them valuable. For instance, the conscious understanding of political and bureaucratic decision-making may be essential to their legitimacy—nothing less will do. Second, the conservative should be comfortable with the notion that there is something preferable about the current system (even we cannot pinpoint exactly what it is), whereby tasks end up automatised through the mediation of the human mind, as compared with a potential system whereby the automation is fully external.

The fourth objection is similar and makes use of the extended mind hypothesis (EMH) to argue that external and internal enhancements are ethically on a par.<sup>59–61</sup> As Levy puts it, "alterations of external props used for thinking are (*ceteris paribus*) ethically on a par with alterations of the brain".<sup>viii</sup> Adopting this, one might argue there is no principled difference between enhancements internal to the human body and externalised automating technologies, provided the combined internal and external system reaches the same outcome.

But the 'ceteris paribus' clause in this principle is all important. There does seem to be a qualitative difference between, say, a calculation performed in the mind and with the help of a calculator. The former seems to be more of a personal achievement than the latter. This points to a general problem with the use of the EMH in this debate. Even if you are persuaded that a functionalist account of the mind lends support to the EMH, you can still acknowledge that there are sufficient, relevant differences between mental processes realised internally (ie, using biological systems) and those that are realised using a combination of internal and external mechanisms. These differences can include the phenomenology of the underlying mental activity, as well as how the activity gets dynamically integrated into the cognitive framework the individual uses to perceive and respond to the world.<sup>62 63</sup> These differences are grist for the conservative mill. We should hold onto values we know attach to predominantly internalised activity, rather than the (potentially different) values that attach to predominantly externalised ones.

The fifth objection tries to argue that my conservative argument is flawed because it ignores the risks and impracticalities that arise from biomedical forms of enhancement. Biomedical enhancement involves interventions that operate on biological systems. But the human body is a complex, evolved system. Tinkering with it through pharmacological or prosthetic enhancements could be risky. Agar puts the point nicely by noting that all proponents of internal biomedical enhancement face an *integration problem*,<sup>14</sup> whereas proponents of externalisation can "cleverly avail themselves of efficiencies enabled by

<sup>viii</sup>Levy<sup>60</sup> quoted in DeMarco and Ford<sup>61</sup> at 318. This is the strong version of the ethical parity principle (EPP). A weaker version states that they are equivalent insofar as the same ethical reasons apply to internal and external forms of enhancement.

biological design” (p. 47). This seems to make externalisation the conservatively superior option.

There are several things to be said in response. First, when Agar describes the integration problem he assumes that the externalisation has an enhancing effect on individual humans. I have already argued that this may not be true. Widespread externalisation could have severe disenchanting effects by making individuals more likely to be worse off relative to the current status quo (eg, through unemployment) or through the atrophy of cognitive capacities. There are, no doubt, risks inherent in the use of biomedical enhancements that do not arise in the case of externalisation. But that does not make externalisation the preferable option. The whole point of the argument I have just defended is that externalisation poses significant risks to our current evaluative equilibrium and it is difficult to mitigate those risks without appealing to some degree of biomedical enhancement. Furthermore, the argument I am making is not that biomedical enhancement should be pursued at all costs. It is simply that biomedical enhancement should be prioritised given its conservative potential.

The final objection is that my argument is unrealistic due to the competing timelines for technological progress. Development of automating technologies is rapid and possibly exponential; development of biomedical enhancements seems to be much slower. We have several putative biomedical enhancers available today, but their effects are rather modest. We could pump more money into their research and development, but there is no guarantee of success. External automating technologies may win the race and we may lose much of our current evaluative equilibrium regardless of our efforts.

My response to this is simple: this may indeed be true, but it only serves to strengthen the argument I am making. If the risk to our evaluative equilibrium really is this great, then evaluative conservatives should be even more wedded to prioritising the proposed solution.<sup>ix</sup>

## CONCLUSION

In conclusion, despite the widespread belief that conservative moral principles are opposed to human enhancement, there are in fact strong reasons to think that human enhancement has conservative potential. This is because technological development does not take place in a vacuum. One cannot consider the effects of biomedical enhancement technology in isolation from other trends in technological progress. When this is done, it becomes apparent that AI, robotics and information technology are developing at a rapid pace and their widespread deployment could undermine much of our current evaluative equilibrium. Biomedical enhancement may be necessary, not merely desirable, if we are to maintain that equilibrium.

<sup>ix</sup>A reviewer worries that the argument implies that individuals—through their decisions to enhance or not—rather than collective institutions bear responsibility for conserving our evaluative equilibrium. They noted that this could be problematic if individuals lacked the means to avail of such interventions or were driven by their social context to avail of them in problematic ways (eg, the poor worker who uses biomedical enhancements to resist bad working conditions but ignores side effects). But this is not a necessary implication of the argument. Although the ultimate *target* of enhancement is the individual, there is no reason why the enhancements themselves cannot be developed and distributed through collective institutions, nor why the individual must bear responsibility for the failure to enhance, or for the negative side effects of enhancement. For a longer discussion, see Danaher.<sup>57</sup>

**Twitter** Follow John Danaher at @JohnDanaher

**Acknowledgements** The author would like to thank Brian Earp and three anonymous reviewers for feedback on earlier drafts of this article.

**Funding** Irish Research Council (New Horizons Grant).

**Competing interests** None declared.

**Provenance and peer review** Not commissioned; internally peer reviewed.

## REFERENCES

- Cohen GA. Rescuing conservatism. In: Wallace RJ, Kumar R, Freeman S, eds. *Reasons and recognition: essays on the philosophy of T.M. Scanlon*. Oxford: OUP, 2011:203–30.
- Cohen GA. Rescuing conservatism: a defense of existing value. In: Otsuka K, ed. *Finding oneself in the other*. Princeton: Princeton University Press, 2012:143–74.
- Pugh J, Kahane G, Savulescu J. Cohen’s conservatism and human enhancement. *J Ethics* 2013;17:331–54.
- Hauskeller M. Human enhancement and the giftedness of life. *Philos Pap* 2011;40:55–79.
- O’Hara K. *Conservatism*. London: Reaktion Books, 2011.
- Lockhart T. *Moral uncertainty and its consequences*. Oxford: OUP, 2000.
- Guerrero A. Don’t know, don’t kill: moral ignorance, culpability, and caution. *Philos Stud* 2007;136:59–97.
- Moller D. Abortion and Moral Risk. *Philosophy* 2011;86:425–43.
- Sepielli A. What to do when you don’t know what to do. In: Shafer-Landau R, ed. *Oxford Studies in Metaethics*. Oxford: OUP, 2009:5–28.
- Buchanan A. *Beyond humanity? The ethics of biomedical enhancement*. Oxford: OUP, 2011.
- Kass L. *Ageless bodies, happy souls: biotechnology and the pursuit of perfection*. The New Atlantis, 2003.
- Fukuyama F. *Our posthuman future*. New York: Farrar, Straus and Giroux, 2002.
- Agar N. *The end of humanity*. Cambridge, MA: MIT Press, 2011.
- Agar N. *Truly human enhancement*. Cambridge, MA: MIT Press, 2013.
- Hauskeller M. *Better humans? Understanding the enhancement project*. London: Acumen, 2013.
- Savulescu J, Sandberg A, Kahane G. Well-being and enhancement. In: Savulescu J, ter Meulen R, Kahane G, eds. *Enhancing human capacities*. Oxford: Blackwell Publishing Ltd, 2011:3–18.
- Kahane G, Savulescu J. The welfarist account of disability. In: Brownlee K, Cureton A, eds. *Disability and disadvantage*. Oxford: Oxford University Press, 2009:15–53.
- Juengst E. What does enhancement mean? In: Parens E, ed. *Enhancing human traits: ethical and social implications*. Georgetown: Georgetown University Press, 1998:29–47.
- Earp BD, Sandberg A, Kahane G, et al. When is diminishment a form of enhancement? Rethinking the enhancement debate in biomedical ethics. *Front Syst Neurosci* 2014;8:12.
- Daniels N. Normal functioning and the treatment-enhancement distinction. *Cambridge Q Healthc Ethics* 2000;9:309–22.
- Harris J. *Enhancing evolution: the ethical case for making better people*. Princeton, NJ: Princeton University Press, 2007.
- Parens E. Is better always good? In: Parens E, ed. *Enhancing human traits*. Washington DC: Georgetown University Press, 1998:1–28.
- Loi M. Technological unemployment and human disenchantment. *Ethics Inf Technol* 2015;17:201–10.
- Persson I, Savulescu J. *Unfit for the future*. Oxford: OUP, 2012.
- Kurzweil R. *The singularity is near: when humans transcend biology*. London: Penguin, 2006.
- Agar N. *The sceptical optimist*. Oxford: OUP, 2015.
- Pratt G. Is a Cambrian explosion coming for robotics? *J Econ Perspect* 2015;29:51–60.
- Ford M. *The rise of the robots: technology and the threat of a jobless future*. New York: Basic Books, 2015.
- Brynjolfsson E, McAfee A. *The second machine age: work, progress, and prosperity in a time of brilliant technologies*. WW Norton and Co, 2014.
- Kaplan J. *Humans need not apply: a guide to wealth and work in the age of artificial intelligence*. Yale University Press, 2015.
- Susskind R, Susskind D. *The future of the professions*. Oxford: OUP, 2015.
- Pasquale F. *The black box society: the secret algorithms and that control money and information*. Harvard: Harvard University Press, 2015.
- Dormehl L. *The formula: how algorithms solve all our problems and create more*. Perigree Books, 2014.
- Lupton D. *The quantified self*. Wiley-Blackwell, 2016.
- Gheaus A, Herzog LM. The goods of work (other than money). *J Soc Philos* 2016. <http://philpapers.org/rec/GHETGO> (accessed 13 Mar 2016).
- Autor DH. Why are there still so many jobs? The history and future of workplace automation. *J Econ Perspect* 2015;29:3–30.

- 37 Srnicek N, Williams A. *Inventing the future: postcapitalism and a world without work*. London: Verso, 2015.
- 38 Hughes J. A Strategic Opening for a Basic Income Guarantee in the Global Crisis Being Created by AI, Robots, Desktop Manufacturing and BioMedicine. *J Evol Tech* 2014;24:45–61.
- 39 Gaus G. *The order of public reason*. Cambridge: Cambridge University Press, 2010.
- 40 Peter F. Political legitimacy. In: Edward NZ, ed. *The Stanford Encyclopedia of Philosophy Spring 2014 Edition*, 2014. <http://plato.stanford.edu/archives/spr2014/entries/legitimacy/>
- 41 Aneesh A. *Virtual migration*. Duke University Press, 2006.
- 42 Aneesh A. Global Labor: Algoratic Modes of Organization. *Social Theory* 2009;27:347–70.
- 43 Danaher J. The threat of algocracy: reality, resistance and accommodation. *Philos Technol* 2016. doi: 10.1007/s13347-015-0211-1, 1/24.
- 44 Zarsky T. Transparent prediction. *Univ Ill Law Rev* 2013;4:1503–70.
- 45 Zarsky T. The trouble with algorithmic decision-making: an analytic road map to examine efficiency and fairness in automated and opaque decision making. *Sci Technol Human Values* 2016;41:118–32.
- 46 Burrell J. How the machine ‘thinks’: understanding opacity in machine learning algorithms. *Big Data Soc* 2016. doi: 10.1177/2053951715622512
- 47 Kitchin R. Thinking critically about and researching algorithms. *Inform Commun Sci* 2016.
- 48 Estlund D. Why not Epistocracy? In: Naomi R, ed. *Desire, identity, and existence: essays in Honour of T.M. Penner*. Academic Printing and Publishing, 2003:53–69.
- 49 Zuboff S. Big other: surveillance capitalism and the prospects of an information civilization. *J Inform Technol* 2015;30:75–89.
- 50 Luper S. Life’s meaning. In: Luper S, ed. *The Cambridge companion to life and death*. Cambridge University Press, 2014:198–214.
- 51 Wolf S. *Meaning in life and why it matters*. Princeton, NJ: Princeton University Press, 2010.
- 52 Carr N. *The glass cage: where automation is taking us*. New York: WW Norton & Co, 2015.
- 53 Van Nimwegen C. et al. The Paradox of the Assisted User: Guidance can be Counterproductive. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* 2006:917–26.
- 54 Burgos D, Nimwegen CV, Oostendorp HV, et al. Game-based learning and the role of feedback. A case study. *Int J Adv Technol Learn* 2007;4: 188–93.
- 55 Dowling C, Leech SA, Moroney R. Audit Support System Design and the Declarative Knowledge of Long-Term Users. *J Emerg Technol Account* 2008;5:99–108.
- 56 Sandel M. *The case against perfection*. Harvard: Belknap, 2009.
- 57 Douglas T. Enhancing moral conformity and enhancing moral worth. *Neuroethics* 2014;7:75–91.
- 58 Danaher J. Human Enhancement, Social Solidarity and the Distribution of Responsibility. *Ethical Theory Moral Pract* 2016;19:359–78.
- 59 Levy N. *Neuroethics: challenges for the 21st century*. Cambridge: Cambridge University Press, 2007.
- 60 Levy N. Rethinking neuroethics in the light of the extended mind thesis. *Am J Bioeth* 2007;7:3–11.
- 61 Levy N. Neuroethics and the Extended Mind. In: Sahakian B, Illes J, eds. *Oxford Handbook of Neuroethics*. Oxford: OUP, 2011:285–94.
- 62 DeMarco J, Ford P. Neuroethics and the Ethical Parity Principle. *Neuroethics* 2014;7:317–25.
- 63 Michaelian K. Is external memory memory? Biological memory and extended mind. *Conscious Cogn* 2012;21:1154–65.