Journal of
APPLIED PHILOSOPHY

Journal of Applied Philosophy, Vol. 40, No. 3, July 2023
doi: 10.1111/japp.12619

Cognitive and Moral Enhancement: A Practical Proposal

EMMA C. GORDON AND VIOLA RAGONESE

ABSTRACT *According to Persson and Savulescu, the risks posed by a morally corrupt minority's potential to abuse cognitive enhancement make it such that we have an urgent imperative to first pursue moral enhancement of humankind – and, consequently, if we are a long way from safe, effective moral enhancement, then we have at least one good reason to consider opposing further cognitive enhancement. However, as Harris points out, such a proposal seems to support delaying life-saving cognitive progress. In this article, we first show that Harris's worry can be expanded to show that Persson and Savulescu's proposal also threatens the development of moral enhancement – precisely what they suggest we have* pro tanto *reason to pursue. From there, we offer our own, alternative proposal – one on which cognitively enhanced researchers play a key role in the production of moral enhancement, and those in the general population who wish to be cognitively enhanced must first accept moral enhancement as an entry requirement. We engage with four substantive objections to our proposal and use these objections to refine and strengthen the details.*

1. Introduction

Persson and Savulescu[1] maintain, controversially, that if we are going to pursue cognitive enhancement, we have an urgent imperative to also pursue a broad moral enhancement of humankind. Although 'enhancement' can refer to any improvement of dispositions and capacities – for example, though education and mild interventions like caffeine – the sort of enhancements Persson and Savulescu have in mind here are actual and possible improvements that take advantage of the latest and future medicine and biotechnology (e.g. pharmaceutical drugs and experimental neural implants) that can take us beyond merely remedying dysfunction in order to make us 'better than well'.[2] In the case of cognitive enhancement specifically, this will potentially include (in addition to traditional interventions like education) interventions such as brain–computer interface technologies and increasingly – as science progresses – fine-grained nootropic drugs that boost focus, memory, and cognitive processing to ever-greater extents. One of the benefits of cognitive enhancement, construed in this strong sense, is that it may help us to more capably address significant and complex dangers faced by humanity, including threats from new diseases, asteroids, and especially climate change.[3]

While cognitive-enhancement proposals have been criticised on other grounds (such as that they are in some way antithetical to human nature, devalue our achievements, or make us less authentic[4]), Persson and Savulescu register an arguably much more serious downside to cognitive enhancement that has received considerable attention in the literature: for Persson and Savulescu, cognitive enhancement is likely to lead (as they put it) to 'ultimate harm'; that is, as they have cautioned, cognitive enhancement carries with it the risk that a (cognitively enhanced) morally corrupt minority will more effectively develop

and deploy nuclear or biological weapons of mass destruction,[5] and that we can expect such a minority to emerge given the size of the human population.[6] An existential threat to humanity's survival is, as Persson and Savulescu see it, relevant and pressing enough to give us a good reason not to pursue cognitive enhancement (despite its acknowledged upsides) unless we are simultaneously pursuing moral bioenhancement,[7] where 'moral enhancement' is meant to pick out the category of (actual and possible) enhancements that are aimed specifically at improving moral capacities and dispositions.[8] In particular, the kind of bioenhancements Persson and Savulescu have in mind include pharmacological bioenhancements targeted at improving altruism and empathy, though they concede that it is unlikely we will see moral bioenhancements that would be effective at the level needed to offset the risks posed by cognitive enhancement unless there are considerable additional investments in scientific research.[9]

This provocative proposal has generated an influx of objections over the past decade. Our intervention in this discussion will attempt to advance the debate in two important ways. First, we offer a hitherto unexplored though serious problem with their proposal, which is that it will be too risky to get off the ground permissibly by their own lights. An appreciation of this problem, however, should not lead us to the bioconservative conclusion that we should forgo moral enhancement. Rather, we want to suggest, there is a way to promote moral enhancement that avoids the kind of risks that would be implied by Persson and Savulescu's proposal. This brings us to the second main payoff of the view advanced here: a new, safer – and practically specific – proposal for moral (as well as cognitive) enhancement that can achieve important benefits with minimal risk.

Here is the plan for the article. Section 2 canvasses a thus far unnoticed objection to Persson and Savulescu's proposal, which takes as a starting point observations by John Harris[10] about specifically cognitive enhancement; we show that this line of critique can be straightforwardly expanded to include a critique of moral enhancement, and in a way that undermines pursuing moral enhancement in the way that Persson and Savulescu suggest. Against the background of this criticism, Section 3 then proposes a practical way to pursue moral enhancement that avoids the kind of problem faced by Persson and Savulescu's proposal; a pleasing benefit of the proposal is that it can be defended alongside pursuing cognitive enhancement – and in a way that Persson and Savulescu's framework would be unable to countenance. Section 4 then considers two potential objections to our proposal, and by engaging with these objections, we suggest some minor improvements and further sharpen the central claims.

2. Costs of Delay: A Problem for Persson and Savulescu

One common criticism of Persson and Savulescu's proposal is that it relies – and implausibly so – on there being no significant overlap between moral and cognitive enhancement[11]; this 'minimal overlap' assumption, after all, seems to be implicit in the very idea that we could pursue one kind of enhancement whilst failing to pursue the other. And yet this assumption is at odds with countenancing cognitive dimensions to moral enhancement (e.g. moral reasoning skills[12]) and ethically relevant dimensions to cognitive enhancement (e.g. notions of intellectual flourishing or wellbeing). Other critiques suggest that substantive moral change also requires changes in the surrounding environment,[13] and in a way that Persson and Savulescu's proposal overlooks.

Additionally, some researchers have taken issue with the substantive details of Persson and Savulescu's prescription of moral enhancement, which focuses specifically on enhancing individuals morally by way of enhancing altruism and a sense of justice.[14]

In this section, we want to raise a new – and, we think, particularly pressing – line of objection, which will take as a starting point concerns raised by Harris[15] pertaining to cognitive enhancement specifically. We will show that a variation of this argument raises trouble for Persson and Savulescu, given that it suggests delaying both cognitive enhancement and moral enhancement.

What Harris observes is that Persson and Savulescu – despite maintaining that cognitive enhancement is unacceptably risky unless moral enhancement is pursued extensively alongside cognitive enhancement – also admit that the development of safer and effective moral enhancement at a level fit for the above task is not a concrete possibility in the immediate future. Harris then notes, rightly, that it seems to follow from the conjunction of their risk-related prescription and their concession about the nascent state of current moral enhancement technologies, that cognitive enhancement should be delayed for a potentially long time, at least as long as it takes to develop more effective and safer moral enhancements. Harris underlines the importance of the costs that delaying cognitive enhancement would have – namely, the 'day-by-day accumulation of premature death and suffering'.[16] For example, currently incurable diseases may continue to kill millions, when they might have been prevented thanks to the innovations cognitive enhancements could afford 'in the clinic and the marketplace that [would] save and ameliorate lives'.[17]

One might worry that Harris's 'costs of delay' argument is framed uncharitably, given that Persson and Savulescu are not claiming that cognitive progress is problematic *per se*, but rather, that the problem is the 'forward momentum' of cognitive enhancement unchecked by moral enhancement.[18] We submit, however, that Harris's delay-based concern is not uncharitable, given their contention that cognitive enhancement has unacceptable risks that can be avoided only by discovering and implementing moral enhancement.[19] We think, accordingly, that the worry that their proposal would entail a significant delay of cognitive enhancement is reasonable and compelling.

Not only do we regard Persson and Savulescu's strategy as provoking the costly delay of cognitive enhancement, but we want to now go further than Harris in suggesting that the same strategy also entails the problematic delay of moral enhancement, and to the extent that this is right, any 'urgent imperative' to pursue moral enhancement through a strategy that (as we are suggesting) has as a consequence a significant delay to moral enhancement, and in such a way as to leave the proposal difficult to permissibly execute by its own lights.

To see why we think such an argument is applicable here, just consider Persson and Savulescu's own admission that 'such a [rapid moral] enhancement could only be effected if significant scientific advances were made'.[20] An internal tension in Persson and Savulescu's proposal appears when we consider (i) the debate-wide consensus that cognitive enhancement may facilitate scientific advances and (ii) Persson and Savulescu's admission that the kind of rapid moral enhancement we need requires 'significant scientific advances'. In short, they purport to give us a reason to delay pursuit of cognitive enhancement until sufficiently safe and effective moral enhancements are available; but – and here is the crux of the internal tension – such a delay on cognitive enhancement comes at the cost of a corresponding delay on moral enhancement, a delay that cognitive enhancement could offset by facilitating

the kinds of scientific advances that for safe and effective moral enhancement are needed. This internal tension in their proposal is not merely a kind of theoretical, philosophical idiosyncrasy, but it also reveals a serious way in which those impressed by Persson and Savulescu's reasoning for the need for urgent moral enhancement face a practical problem. In short, the very kinds of existential risks that lead champions of Persson and Savulescu's proposal to embrace rapid moral enhancement are risks that – for all Persson and Savulescu have said – would need to be initially incurred in order to expedite the kind of urgent scientific advances that are needed to make possible the kind of safe and effective moral enhancement that is needed to combat these very risks.

What we are left with – by way of metaphor – is a kind of 'spinning wheel of delays' – viz., a situation where genuine and concerning threats require moral enhancement, but as soon as we want to take concrete steps to pursue the latter, we are faced with a series of delays that make it difficult to locate any practical starting point. In this respect, the argument we have sketched exposes an important sense in which Persson and Savulescu's proposal is impractical; their prescriptions, along with their concessions about the current state of science, leave us ill equipped to achieve the results they have prescribed.

3. A New Proposal for Safe Cognitive and Moral Enhancement

The situation for those impressed by Persson and Savulescu's prescription for urgent moral enhancement is not as bleak, we want to now suggest, as things might appear. Rather than to either downplay the risks posed by cognitive enhancement or to caution (as bioconservative philosophers have) against moral enhancement, we want to instead outline the core elements of a strategy for pursuing moral enhancement that is both practical and – perhaps surprisingly – can be pursued safely alongside cognitive enhancement.

In order to get the kind of result we want, we are going to outline three specific desiderata that we think – informed by the discussion in Section 2 – that a practical and safe proposal for moral and cognitive enhancement should meet:

> (**Desideratum 1:** threat concession). The proposal must take full cogniance of the existential dangers Persson and Savulescu highlight and which unchecked cognitive enhancement would stand to pose.
> (**Desideratum 2:** acknowledge cost of delaying cognitive progress). The proposal must be sensitive to worries about the cost of delaying cognitive progress.
> (**Desideratum 3:** avoid unnecessary delay to moral enhancement). The proposal must accommodate costs of delaying of moral enhancement (as suggested in Section 2).

With reference to Desiderata 1–3, we can now appreciate that Persson and Savulescu's proposal meets Desideratum 1 while failing both Desiderata 2 and 3. We want to now outline a proposal that can do better. We take as a starting point an observation key to Desideratum 3, which is that significant scientific advancements would be key to the production of moral enhancement. Of course, Persson and Savulescu caution that pursuing cognitive enhancement (prior to safe and effective moral enhancement) exposes us to existential risks generated by an immoral minority of cognitively enhanced agents.

It is in this area that we see logical space for a very specific type of proposal which – as far as we see it – is the only viable strategy that can meet all of Desiderata 1–3. The central idea of the proposal is this: as more sophisticated forms of cognitive enhancement become accessible, they should be made available in a carefully regulated way to researchers whose aim is to find and develop safe and effective moral enhancement. Consider that success in medical research requires high cognitive skills; concentration, cognitive flexibility, memory, executive function in complex cognitive tasks. Cognitive enhancement of researchers specifically, as opposed to the general population (which carries risks that we countenance in Desideratum 1), minimises the risk of immoral use, while at the same time amplifying and facilitating the cognition of those whose cognition is most urgently needed to make breakthroughs in the science of moral enhancement. Even more, we suggest, the specific kinds of cognitive enhancements to which we should direct research funding and promote legislation to allow (in a regulated fashion) are cognitive enhancements most likely to facilitate the cognitive progress in moral enhancement specifically.

In more concrete terms, one starting point will be cognitive enhancement aimed at facilitating the work of research into pharmaceutical-based moral enhancement, where advances in research on such pharmaceuticals is suggestive of an increasing capacity to improve 'moral decision making and morally significant behaviour'.[21] However, in practice, targeting moral-enhancement researchers for cognitive enhancement is not limited to researchers who focus on pharmaceutically driven moral enhancement. Perhaps even more effective forms of moral enhancement will involve additional research into morally relevant biotechnology (e.g. brain–computer interfaces, Socratic AI,[22] etc.). The idea in principle remains the same: the relevant focus group for cognitive enhancement should be, in the first instance, those most likely to help bring about the moral enhancement, and given the current scientific barriers to developing safe and effective moral enhancement, this group will be scientific researchers invested in the production of new and improved moral enhancements.

The strategy described above manages to hit all three of Desiderata 1–3. It satisfies Desideratum 1 (threat concession) in that it places a premium on developing moral enhancement while also significantly curtailing cognitive enhancement of the sort that would engender additional risk. It satisfies Desideratum 2 (acknowledge cost of delaying cognitive progress) in that (unlike what Persson and Savulescu have proposed) it prescribes an important dimension of cognitive progress, and it does so through promoting (in a restricted way) a targeted form of cognitive enhancement. And crucially, the proposal satisfies Desideratum 3 (avoid unnecessary delay to moral enhancement), and it does this by facilitating exactly what would be needed to most expediently meet a thus far unsatisfied necessary condition on safe and effective moral enhancement, which is additional scientific research into such enhancements that by Persson and Savulescu's admission is currently lacking. Although directing resources towards the specific policy of promoting and making available cognitive enhancement for researchers aiming to improve the science of moral enhancement is a very limited and specific proposal, the fact that it has the capacity to satisfy all three desiderata above – and thus navigate a way out of the kind of impasse of delays discussed in Section 2 – is enough to make the proposal an interesting one that is deserving of theoretical and practical attention.

A critic might quibble with us on the point of whether the proposal – on closer inspection – satisfies Desideratum 2. Perhaps – as the anticipated worry would go – satisfactorily acknowledging the costs of delaying cognitive progress will require going

further than we have to address (and aim to offset) costs that forgoing cognitive enhancement will have for the general population not engaged in research into moral enhancement. For all we have said so far, for example, scientific progress in other areas of human interest and concern will be delayed precisely because resources invested in cognitive enhancement will be diverted specifically to moral-enhancement research. How, then, does our proposal acknowledge these losses?

In response, we want to first register that our position, in order to countenance Desideratum 2 without at the same time flying in the face of Desideratum 1, requires that we do not go too far to meet Desideratum 2. There is, as we see it, no clear way to offset all costs associated with forgoing cognitive enhancement without falling foul of Desideratum 1, which requires satisfactorily acknowledging the risks of cognitive enhancement in a general population.

That being said, we do think we can go further than we have thus far to explain how we meet Desideratum 2 without compromising the way we have met Desideratum 1. This will involve considering in a bit more detail how the proposal can be theoretically organised in two stages. First, there is the research-driven stage – which we have described – where cognitive enhancement is prioritised for those researchers invested in developing safe and effective moral enhancement that is not yet available,[23] where – crucially – the necessity of targeting moral-enhancement researchers (for the dissemination of regulated cognitive enhancement)[24] is predicated on the moral science being in its nascency. Persson and Savulescu might be right that presently we lack the science to deliver safe and effective moral enhancements; however, if our proposal described succeeds in its aims, then this will be temporary. There will be a point at which, through investment in enhancing priority research in moral enhancement, we do have the capacity to disseminate moral-enhancement options (e.g. pharmacological and/or biotechnological) in a population that are safe and suitably effective.

At this future point, we can envision a 'second stage' of the proposal that becomes relevant – a stage triggered if and only if moral enhancement is developed that is safe and effective. At this point, we suggest a new balance in policy is needed to continue to meet Desiderata 1–3 (one, we think, that will ultimately allow us to go much further in satisfying Desideratum 2 without risking failing to satisfy Desideratum 1). During this second stage, moral enhancement (which is at this point safe and effective, *ex hypothesi*) will be available widely in a population for those who wish to take it. Even more, moral enhancement should be encouraged – even for those who are not initially interested in being cognitively enhanced. That is because, as noted above, there is good cause to think that we are already morally defective, regardless of how future cognitive enhancement might affect us. Nonetheless, we submit that moral enhancement should not be compulsory, because – as Giubilini and Minerva suggest, 'imposing any medical intervention would impermissibly violate bodily autonomy'.[25]

That said, and here is where we speak directly to Desideratum 2, we propose that – at this second stage of the proposal – cognitive enhancement should be available in a population (and in some cases, encouraged) although only to those who have already voluntarily availed themselves of moral enhancement. This qualification is important. It implies that we will not get cases of the sort that Desideratum 1 identifies as risky: cases where cognitive enhancements are used by an immoral minority of the population to disastrous ends. Our proposal safeguards against this by permitting such cognitive enhancement only for those who volunteer to accept the latest and best moral enhancement. An

unintended benefit (vis-à-vis Desideratum 3) of this policy is that it will de facto incentivise individuals who might otherwise not be interested in voluntary moral enhancement to elect to undergo such enhancement, given that doing so will be a necessary precondition for availing themselves of the kinds of cognitive enhancements that can help them pursue individual goals more expediently. Finally, and crucially, on the proposal (at this second stage) under consideration, we are significantly widening the scope of the cognitively enhanced, relative to what we have proposed at the first stage. And in this way, while the first stage of our proposal goes some distance to countenancing Desideratum 2 (i.e. by promoting cognitive enhancement in a select group of researchers), we go even further at the second stage, where such cognitive enhancement will be less risky, by permitting it in a wider population under conditions where that population is voluntarily morally enhanced.

4. Objections and Replies

We want to now consider in some detail four anticipated objections to our proposal, taking into account both stages as described in Section 3. The first two concern whether our problem fails Desideratum 1 at the first stage, even if not at the second. The third is a more general criticism of the proposal in the light of reflections on the distributed mechanisms of scientific-knowledge generation. The fourth engages with concerns related to enhancement, including moral enhancement, and inequality. We take these objections in turn.

4.1. Objection 1

The first stage of the proposal countenances the provision of cognitive enhancement to a group of researchers; however, these researchers will (*de facto*) not yet be morally enhanced themselves. Researchers, like any subsector of the population, will presumably include a distribution of individuals across the spectrum of morality. Some researchers, presumably even those who are engaged in developing moral enhancements, will be morally corrupt; they might have ulterior motives beyond moral-enhancement research. And so there is an inevitable risk that such researchers, equipped with cognitive enhancement although without corresponding moral enhancement, will provoke harm and destruction in precisely the way feared by Persson and Savulescu. Thus, as the line of thought goes, our proposal – at the first stage – invites a kind of 'loophole' whereby the proposal does not satisfactorily countenance Desideratum 1. It does not suitably take into account the risks that cognitive enhancements will fall into the hands of an immoral minority.

We want to offer two reasons to think the above 'evil researcher' worry is not as serious as it might initially look. The first reason follows from Persson and Savulescu claiming that 'in a huge human population, there are bound to be individuals that are depraved enough to want to destroy us all'.[26] We agree with them on this point, but we want to stress that the probability of such individuals in the human population at large is considerably different from the probability of such individuals in a small, carefully selected group of scientists than in the entire human population. To be clear, we concede entirely that there will be morally imperfect individuals, perhaps significantly so, within the carefully selected group of individuals who will comprise the selection of those moral-enhancement researchers selected for cognitive enhancement. We emphasise, however, that the threat Persson

and Savulescu highlight is not best captured as materialising when merely morally bad individuals are cognitively enhanced. The relevant risk event here is likely to be the use of increased intelligence for, e.g. financially motivated corruption. The kind of existential threat that drives Persson and Savulescu's own proposal – and which we are aiming to address in satisfying Desideratum 1 – is a threat that individuals with an interest in, e.g. developing technologies aimed at destroying the planet or our species will be capabilised through cognitive enhancement. However, to reiterate, the likelihood of those such individuals within a select and relatively small group of researchers is negligible compared to the (already small) but not insignificant proportion of such individuals in the general population. Granted, one might object that we have not established that the probability is zero. We concede this; however, we do not take it that satisfactorily addressing Desideratum 1 requires taking the probability down to zero (just as, by way of comparison, we do not think satisfying Desideratum 2 requires avoiding all possible costs of foregoing enhancement in the general population).

Secondly, we want to emphasise an additional kind of control in place to help minimise the already very low risk that the cognitively enhanced group of researchers would include individuals who (despite already researching moral enhancement) have desires that present existential threats and their acting on would then be made possible through cognitive enhancement. The control, we suggest, will include significant psychological profiling of all those among the moral-enhancement researchers who would be candidates for cognitive enhancement. Once again – and in connection with the previous point – we submit that the requisite psychological profiling would not be aiming to individuate people who are merely morally flawed or otherwise. The screening will be instead for individuals who score high enough on metrics associated with being deeply problematic; the question of which metrics would be prioritised should be outsourced to psychological as opposed to moral philosophers.

Taken together, our responses to the initial worry, we think, suitably offset the kind of argument envisaged that would maintain that we have not suitably met Desideratum 1 (at least, at the first stage of the proposal).

4.2. Objection 2

Here is a second objection. Even if we grant – as per our response to Objection 1 – that as there is a nonzero probability of individuals intent on mass destruction and who are also within the class of moral-enhancement researchers afforded cognitive enhancement, a separate kind of conditional probability also deserves serious attention. Namely, this separate concern is that if such cognitive enhancement is pursued to the extent that it can be distributed to the core team of moral-enhancement researchers (during stage 1 of the proposal), then, it seems, the probability that these cognitive enhancements then fall into the hands of the general public (e.g. through accident, technical or human error) will accordingly be raised significantly relative to a situation where such cognitive enhancements are not even developed or pursued under the restricted conditions (for moral-enhancement research teams) we advocate. And then, on the supposition that our proposal runs this risk of letting new cognitive enhancements fall into the general population, the probability of mass destruction resulting from cognitively enhanced wicked individuals intent on mass destruction rises[27] (relative to what it would be on a proposal that did not even promote cognitive enhancement in the restrictive circumstances we encourage).

Our response to the above objection is not to dispute the existence of the kind of risk identified, but rather, to maintain that such a risk can reasonably be set aside as not (nearly) significant enough to outweigh the risks – including existential risks – we plausibly run by failing to develop, promote, and disseminate new and improved moral enhancement. In short, we cannot (right now, at least) have it all. We lack, as Persson and Savulescu note, the science presently to facilitate safe and effective moral enhancement that are equal to the task of helping us to offset existential risks we presently face. The already present risks must be taken into account when considering the threshold of risk to tolerate when asking whether a moral-enhancement proposal satisfactorily meets Desideratum 1.

Here it is worth reiterating that – in keeping with our response to Objection 1 – we do not take it that satisfactorily countenancing Desideratum 1 – and thus taking seriously the threats posed by unchecked cognitive enhancement in a population – requires reducing the objective probability of the relevant risk event (in this case, mass destruction via cognitively enhanced individuals with destructive intent) to zero. Nor, for that matter, do we maintain that the probability of the relevant risk should be low enough to be *de minimis* and rightly ignored. Addressing the dilemma raised by Persson and Savulescu requires, from a philosophical standpoint, an assessment of costs and benefits, against a background situation that is already a vexed one.

4.3. Objection 3

A third kind of objection we might envision targets what appears to be an overly individualised and mistaken characterisation of scientific progress that our proposal would seem to imply.

A prevalent idea at the intersection of social epistemology, the philosophy of science and of distributed cognition, stresses the collaborative nature of scientific discoveries.[28] Take, for example, the case of scientific research teams, particularly in the case of large-scale scientific breakthroughs. When such researchers generate new scientific knowledge, this is rarely a product of any one individual's high intelligence alone. Rather – think here, for illustration, of the discovery of gravitational waves in 2015[29] – the new scientific knowledge is generated by many individual thinkers working together. And, what is more, as, e.g. de Ridder[30] and Palermos[31] point out, often times the resulting scientific knowledge is best understood as a discovery creditable to the interactions of the group as a whole, given that often enough discrete tasks are distributed on the basis of individual-level expertise across a scientific research team, which then interacts in ways that reliably generate an overall result.[32]

With this in mind, it might seem as though the very idea of cognitively enhancing researchers moral enhancement will be underdetermining of any tangible results, given that the individual scientists aren't likely in practice to achieve any notable discovery through individualised cognition which would be affected through moral enhancement.

In response to this kind of concern, we want to emphasise that while our proposal involves individual-level cognitive enhancement, it is in principle compatible with various ways in which scientific collaborations amongst researchers could be realised. This, we should stress, includes compatibility (as it should) with the foregoing view that discoveries are often collective efforts.

We want to explain this compatibility in two ways. First, consider that even in paradigmatic cases of distributed cognitive tasks pursued by a scientific research team, the basis

on which individual tasks might be most efficiently allocated within such a team is individual-level expertise.[33] Increases to the capacity of individuals to fulfill individually designated tasks within a wider team will plausibly (on distributed models, e.g. such as Palermos'[34]) result in a more effective collective result even when we hold fixed the distributed assumption that the result attained by the group is principally creditable to the collective.

Secondly, various targets of cognitive enhancement (e.g. memory, focus) offer the capacity to boost cognitive performance not only individual-level thinking tasks, but also when such tasks are repurposed to facilitate effective collaboration on a distributed task. To use just memory and focus alone as examples, consider that both of these features of one's enhanced intelligence will be effective in more skillful interactions with other members on a research team, and this includes the task of quality monitoring (see, e.g. Palermos[35]) as a form of 'checks and balances' on other team members' contributions to the whole.

In sum, then, while cognitively individual-level researchers looks ostensibly like it would fit most naturally with a conception of scientific progress that is implausibly individualised, this needn't be so; the proposal fits well with the prevailing idea that our best research results in science come as a result of effective intellectual collaborations.[36]

4.4. Objection 4

There have been long-standing worries in the enhancement literature that promoting enhancement risks exacerbating existing inequalities. Consider, for instance, how Bostrom and Roache express this overarching worry in the case of cognitive enhancement:

> people with radically enhanced cognitive capacities might gain vast advantages in terms of income, strategic planning, and the ability to influence others; in other words, an enhanced cognitive elite may gain socially significant amounts of power.[37]

One might initially think that the above kind of concern, even if valid as a concern about cognitive enhancement, is inapplicable to specifically moral enhancement – given that, by gaining moral enhancement, one will presumably then be less likely to exercise whatever power one has over those in worse-off positions. However, as Robert Sparrow has noted,[38] programmes that promote moral enhancement also generate their own distinctive kinds of inequality-based risks. Two such examples Sparrow mentions are, first, that the very idea that we can reliably identify individuals on the basis of being morally better or worse – something he takes to be implicated by any society-wide moral enhancement programme – is in conflict with egalitarian ideals.[39] Secondly, Sparrow raises the worry that in a society that promote moral enhancement, there is a real risk that the morally enhanced will *de facto* be overrepresented in political decision-making.[40]

Against the above background, we can envisage an inequality-based strand of objection to the proposal developed here with two components to it. First, at stage 1 of the proposal, there is a risk that cognitive enhancements will not be distributed just for the purposes of moral-enhancement research, and in a way that invites Bostrom and Roache's concern about widening power gaps between the cognitively enhanced and unenhanced. Second, and perhaps more critically, the end result of the proposal at stage 2, in which moral enhancement is made available, might then be subject to Sparrow's worries. We have both

a concession and several replies to this pair of equality-based worries. The concession is that neither risk can be entirely eliminated.[41] This concession is not distinctly applicable to our proposal, which is one (as we will emphasise) that can mitigate against both risks. But the concession of such risks is simply a fact of any enhancement proposal, and the important question in the face of that concession is whether the risks can be sufficiently mitigated against, and secondly, whether whatever residual risks to inequality that are present even after attempts to mitigate against them are greater or less great than risks of forgoing the enhancement proposal altogether.

We want to address first the mitigation point. Our proposal at stage 1 is not as susceptible to the Bostrom–Roache-style objection as more generic cognitive-enhancement proposals because our proposal restricts the availability of such enhancements to a minority. Our proposal is mitigating against Sparrow-style risks at stage 2, for the reason that moral enhancements on the proposal will only be available to volunteers. This proviso is aimed not only at insulating the proposal from objections to, e.g. state-mandated moral enhancement, but also to inequality-based worries that would arise were moral enhancement only minimally available.

Even so, as we concede, the above mitigations do not eliminate equality-based risks entirely. However, and this is the second key point of the response, we think the proposal on the whole does well when we ask whether whatever residual risks to inequality that are present even after attempts to mitigate against them are greater or less great than risks of forgoing the enhancement proposal altogether. Here, we invite the reader to contextualise the proposal as a response to the predicament Persson and Savulescu begin with – one on which our present trend would (as they see it) represent an increasing existential risk. Within this context, it is far from clear that residual risks to inequality that would arise were the programme imperfectly implemented[42] would be more significant than those Persson and Savulse have antecedently identified.

5. Concluding Remarks

Despite delivering a controversial positive programme for moral enhancement, the fact that Persson and Savulescu's proposal has received as much attention as it has is, we think, a testament to the fact that they have put their finger on a very serious risk with interesting philosophical (and practical) significance. Moral philosophers and bioethicists have wrestled with the matter of how (with or without moral enhancement) to best address the problem Persson and Savulescu raise and try (we think unsuccessfully) to address themselves.

This article has tried to take two important steps forward in advancing the debate surrounding these issues. First, we have identified what we take to be an overlooked but very serious problem that faces Persson and Savulescu's proposal. The problem, suitably appreciated, makes the prospects of progress bleak for addressing the original existential-risk-related problem Persson and Savulescu have identified. We use this problem for Persson and Savulescu as a lens through which to articulate three desiderata that any answer to the problem they have identified should meet. Any such proposal must (i) countenance the kinds of dangers that unchecked cognitive enhancement poses; (ii) be suitably sensitive to worries about the cost of delaying cognitive progress; and (iii) accommodate the costs of delaying moral enhancement. Whereas Persson and Savulescu's own proposal fail the second and third of these desiderata, we have offered

a practical proposal that, to our knowledge, offers the only way to strike a balance so as to satisfactorily accommodate each of the above three desiderata. Having developed this proposal – which focuses on highly selective cognitive enhancement aimed at facilitating research into moral enhancement – we have considered several objections to the proposal, and we have argued that the kind of recommendations we have outlined can withstand each of them.

Emma C. Gordon, University of Glasgow, Glasgow, UK. emma.gordon@glasgow.ac.uk

Viola Ragonese, University of St Andrews, St Andrews, UK

Acknowledgements

The authors would like to think anonymous referees at the *Journal of Applied Philosophy* for thoughtful and detailed feedback that improved this article. Gordon's work on this article was conducted as part of the Leverhulme-funded 'A Virtue Epistemology of Trust' (#RPG-2019-302) project, which is hosted by the University of Glasgow's COGITO Epistemology Research Centre, and she is grateful to the Leverhulme Trust for supporting this research.

NOTES

1 Persson, Ingmar, and Julian Savulescu. 2008. "The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity." *Journal of Applied Philosophy* 25(3): 162–77; Persson, Ingmar, and Julien Savulescu. 2015. "Summary of Unfit for Future." *Journal of Medical Ethics* 41(3): 338.

2 Bostrom, Nick, and Anders Sandberg. 2009. "Cognitive Enhancement: Methods, Ethics, Regulatory Challenges." *Science and Engineering Ethics* 15 (3): 311–41.

3 See e.g. Harris, John. 2011. "Moral Enhancement and Freedom." *Bioethics* 25(2): 102–11.

4 For some prominent criticisms about cognitive enhancement's impact on human nature and human dignity, see Kass, Leon. 2003. "Beyond Therapy: Biotechnology and the Pursuit of Human Improvement." In *President's Council on Bioethics*. Washington, DC; Kass, Leon. 2004. *Life, Liberty and the Defense of Dignity: The Challenge for Bioethics*. San Francisco: Encounter Books. For discussion of whether cognitive enhancement makes us less authentic, see e.g. Pugh, Jonathan, Hannah Maslen, and Julian Savulescu. 2017. "Deep Brain Stimulation, Authenticity and Value." *Cambridge Quarterly of Healthcare Ethics* 26(4): 640–57; Maslen, Hannah, and Julian Savulescu. 2014. "Pharmacological Cognitive Enhancement – How Neuroscientific Research Could Advance Ethical Debate." *Frontiers and Systems Neuroscience* 8: 107.

5 Persson and Savulescu 2008 op. cit., p. 166.

6 Persson and Savulescu 2008 op. cit., p. 174.

7 Persson and Savulescu 2015 op. cit., p. 338. They regard traditional moral enhancement (e.g. moral education) as ineffective since improving morally seems harder than improving cognitively, perhaps because it requires motivation to act morally on top of knowledge of what is right. Instead, they suggest we focus on developing drugs that enhance altruism and our sense of justice – Persson and Savulescu take these two aspects to be central to morality (2008 op. cit., pp. 168–9) and surmise that their proposal is a fairly uncontentious one that people who endorse a wide variety of normative ethical theories should support. Persson and Savulescu's earlier work (2008 op. cit., p. 174) suggests that safe and effective moral enhancement should be compulsory, but they abandon this position later (e.g. 2015 op. cit.).

8 For a notable line of critique in the special case of characterising 'moral enhancement' that is both practical and philosophically interesting, see Earp, Brian D. 2018. "Psychedelic Moral Enhancement." *Royal Institute of Philosophy Supplement* 83: 415–39. For a survey of some different definitional approaches to both cognitive and moral enhancement, see Gyngell, Chris, and Michael J. Selgelid. 2016. "Human Enhancement:

Conceptual Clarity and Moral Significance." In *The Ethics of Human Enhancement*, edited by Steve Clarke, Julian Savulescu, C.A.J. Coady, Alberto Giubilini, and Sagar Sanyal, 111–26. Oxford: Oxford University Press.

9 Persson and Savulescu 2008 op. cit., pp. 172–4.

10 Harris 2011 op. cit.

11 For arguments that moral enhancement involves cognitive enhancement and cognitive enhancement involves moral enhancement, see e.g. Harris 2011 op. cit. Adam, Carter, J., and Emma C. Gordon. 2015. "On Cognitive and Moral Enhancement: A Reply to Savulescu and Persson." *Bioethics* 29(3): 153–61. We will for present purposes set aside these sorts of worries and embrace cognitive-moral enhancement dichotomy endorsed by Persson and Savulescu.

12 There is an emerging consensus for supposing a cognitive approach to moral bioenhancement will be promising, as noted by Shaw, Elizabeth. 2018. "Retributivism and the Moral Enhancement of Criminals Through Brain Interventions." In *Moral Enhancement: Critical Perspectives*, edited by Michael Hauskeller and Lewis Coyne, 251–70. Cambridge: Oxford University Press. For a few examples, see e.g. Schaefer, G. Owen, and Julian Savulescu. 2016. "Procedural Moral Enhancement." *Neuroethics* 12: 73–84; Paulo, Norbert, and Jan Christoph Bublitz. 2017. "How (Not) to Argue for Moral Enhancement: Reflections on a Decade of Debate." *Topoi* 38: 95–109. The former advocate enhancing the cognitive capacities that would enable one to more reliably make correct moral judgments while the latter stress the importance of rational capacities in exercising moral agency effectively in a democracy. Likewise, see e.g. Earp, Brian, Thomas Douglas, and Julian Savulescu. 2017. "Moral Neuroenhancement." In *The Routledge Handbook of Neuroethics*, edited by L. Syd M. Johnson and Karen S. Rommelfanger, 111–26. New York: Routledge. Here, Earp *et al.* stress the importance of cognition in moral enhancement by noting the importance of higher-order capacities to modulate one's moral responses in a flexible, reason-sensitive, and context-dependent way (p. 171). See also Shaw 2018 op. cit. sect. 2.2 for an overview.

13 See e.g. Hardcastle, Valerie. 2018. "Lone Wolf Terrorists and the Impotence of Moral Enhancement." *Royal Institute of Philosophy Supplement* 83: 271–91 for a case study of self-radicalised individuals that illustrates the importance of environmental factors in determining moral behaviour (on the basis of which she then argues that moral bioenhancement would be less effective than its supporters suggest).

14 See e.g. de Melo-Martín, Inmaculada. 2018. "The Trouble with Moral Enhancement." *Royal Institute of Philosophy Supplement* 83: 19–33 for why we might need to worry that those who have their altruism and sense of justice enhanced will be more likely to commit violence in response to perceived injustice. For further critical responses to Persson and Savulescu, see Jotterand, Fabrice, and Susan Levin. 2017. "Moral Deficits, Moral Motivation and the Feasibility of Moral Bioenhancement." *Topoi* 38(7): 63–71; Azevedo, Marco A. 2016. "The Misfortunes of Moral Enhancement." *Journal of Medicine and Philosophy* 41(5): 461–97.

15 Harris 2011 op. cit.

16 Harris 2011 op. cit., p. 109.

17 Harris 2011 op. cit., p. 109.

18 Persson and Savulescu 2008 op. cit., p. 162.

19 Persson and Savulescu 2008 op. cit., p. 174.

20 Persson and Savulescu 2008 op. cit., p. 173.

21 Levy, Neil, Thomas Douglas, Guy Kahane, Sylvia Terbeck, Philip J. Cowen, Miles Hewstone, and Julian Savulescu. 2014. "Are You Morally Modified?: The Moral Effects of Widely Used Pharmaceuticals." *Philosophy, Psychiatry & Psychology* 21(2): 111–25. See also Earp 2018 op. cit. for a persuasive proposal that using psychedelics as an adjunct to moral enhancement would be more effective than other, more direct interventions. For exploration of propranolol's impact on racial bias, see also DeGrazia, David. 2014. "Moral Enhancement, Freedom and What We (Should) Value in Moral Behavior." *Journal of Medical Ethics* 40(6): 361–68. Meanwhile, for how oxytocin might promote love and trust, see e.g. Zak, Paul, Robert Kurzban, and William T. Matzner. 2014. "The Neurobiology of Trust." *Annals of the New York Academy of Sciences* 1032: 224–7. That said, it remains contested whether and to what extent the reported results give us reason to find (at least present) medications to be *all-things-considered* moral enhancers. For example, Agar, Nicholas. 2015. "Moral Bioenhancement is Dangerous." *Journal of Medical Ethics* 41: 343–45 notes some existing strategies for morally enhancing empathy might not be suitably fine-grained. For criticism of SSRIs in particular, see Wiseman, Harris. 2014. "SSRIs as Moral Enhancement Interventions: A Practical Dead End." *AJOB Neuroscience* 5(3): 21–30.

22 See, for example, Lara, Franciso, and Jan Deckers. 2020. "Artificial Intelligence as a Socratic Assistant for Moral Enhancement." *Neuroethics* 13(3): 275–87.

23 It is worth registering that in principle we see no incompatibility with the proposal outlined here at the first stage and the desirability of open science. Research results and methods generated by (cognitively enhanced) researchers on moral enhancement can compatibly, given what is recommended on the present proposal, be distributed outside the scientific-community via open access. This includes, it should be emphasised, full disclosure of methods. What is restricted is access to the cognitive enhancements themselves, not information generated by the research teams.

24 As an anonymous reviewer notes, the development process of moral enhancement, as we conceive of it, could potentially be accelerated if cognitive enhancement would be provided to all researchers involved in developing the final product 'moral enhancer', as well as cognitive enhancers. Such acceleration of the process would, we think, have a straightforward benefit; there is a worry that this accelerated proposal would generate a variation of the 'immoral researcher' objection; however, we take it that our response to that objection already is applicable *mutatis mutandis* to the accelerated proposal. Thanks to the anonymous reviewer for noting this potential variation on the proposal.

25 Giubilini, Alberto, and Francesca Minerva. 2019. "Enhancing Equality." *Journal of Medicine and Philosophy* 44(3): 335–54, p. 352.

26 Persson and Savulescu 2008 op. cit., p. 174.

27 That is, it rises *conditioned upon* the rising probability (afforded by our proposal) that cognitive enhancements aimed to be distributed exclusively to moral-enhancement researchers fall by accident into circulation in the general public.

28 See e.g. Hardwig, John. 1985. "Epistemic Dependence." *Journal of Philosophy* 82(7): 335–49; Giere, Ronald. 2002. "Scientific Cognition as Distributed Cognition." In *Cognitive Bases of Science*, edited by Peter Carruthers, Stephen Stich, and Michael Siegal. Cambridge: Cambridge University Press; De Ridder, Jeroen. 2014. "Epistemic Dependence and Collective Scientific Knowledge." *Synthese* 191(1): 37–53, Palermos, S. Orestis. 2020. "Epistemic Collaborations: Distributed Cognition and Virtue Reliabilism." *Erkenntnis* 87: 1481–500.

29 For an overview of this discovery, see Collins, Harry. 2017. *Gravity's Kiss: The Detection of Gravitational Waves*. Cambridge, MA: MIT Press.

30 De Ridder 2014 op. cit.

31 Palermos 2020 op. cit.

32 For discussion of how collective cognition is often not reducible to individual-level knowledge when a task is distributed across a system, see e.g. Hutchins, Edwin. 1995. *Cognition in the Wild*. Cambridge, MA: MIT Press.

33 See De Ridder 2014 op. cit., p. 46.

34 Palermos 2020 op. cit.

35 Palermos 2020 op. cit.

36 Thanks to a reviewer at the *Journal of Applied Philosophy* for recommending further discussion on this point.

37 Bostrom, Nick, and Rebecca Roache. 2007. "Ethical Issues in Human Enhancement." In *New Waves in Applied Ethics*, edited by Jesper Ryberg, Thomas S. Petersen, and Clark Wolf, 120–52. Basingstoke: Palgrave Macmillan.

38 Sparrow, Robert. 2014. "Egalitarianism and Moral Bioenhancement." *American Journal of Bioethics* 14(4): 20–8.

39 For discussion more generally on this point, see Giubilini and Minerva 2019 op. cit.

40 Sparrow 2014 op. cit.

41 Note that we are granting this point even though the proposal itself involves making moral enhancement available unrestrictedly.

42 This includes the initial inequality that features in access to cognitive enhancement that is applicable to researchers developing moral enhancement.