

Published as: Carter S. (2015) 'Putting a price on empathy: against incentivising moral enhancement', *Journal of Medical Ethics*, 41(10); 825–829.

Putting a Price on Empathy: Against Incentivising Moral Enhancement

The debate about moral bioenhancementⁱ began with Persson and Savulescu's[1] claim that scientific progress accelerated by cognitive enhancement would lead to possession of a level of power with which our under-evolved moral psyche would be unable to cope, leading us to an instance of ultimate harm.

They asserted that the only solution to avoid such a disaster would be to take up a programme of compulsory moral enhancement alongside cognitive enhancement[1]. Recently, however, Rakic[2] has argued that the possibility of self-annihilation, of such instances of ultimate harm, can never be eliminated even with the use of compulsory moral enhancement. Instead, Rakic states that we can only aim at keeping the likelihood of ultimate harm at a *minimum* (rather than eradicating the risk altogether), which he notes seems insufficient to justify the use of compulsory moral enhancement. Instead, he proposes a programme of voluntary moral enhancement that would be incentivised in order to encourage participation.

This article aims to develop the debate by exploring the idea of incentivising programmes of voluntary moral enhancement. I will demonstrate that, while it is unclear that areas of concern more typically raised in response to incentives (such as coercion and social justice) are necessarily relevant in this instance, public attitudes relating to the characteristics likely to be affected by moral enhancementⁱⁱ suggest that any proposal to incentivise moral enhancement could be met by, at best, public indifference to the idea and at worst, moral outrage.

It is beyond the scope of this paper to engage directly with the ongoing debates surrounding moral enhancement generally; my concern here is with the more specific area of the ethics and efficacy of incentivised programmes of moral enhancement. This article seeks to demonstrate that, regardless of one's view of moral enhancement itself, any plan to incentivise a voluntary programme of the endeavour would prove both practically and ethically problematic.

The Argument for Incentives

One problem in particular becomes immediately obvious when considering any programme of voluntary moral enhancement: namely, it seems reasonable to assume that very few people would be likely to volunteer themselves to undergo the intervention. This could be because most people would not see the point of undergoing moral enhancement as they would not consider themselves

ⁱ Hereafter bioenhancement is referred to simply as 'enhancement'

ⁱⁱ As reported in[3]

to need it. Relatedly, people who might be seen to most need the intervention could be the least likely to volunteerⁱⁱⁱ. As Persson and Savulescu note:

If safe moral enhancements are ever developed, there are strong reasons to believe that their use should be obligatory, like education or fluoride in the water, **since those who should take them are least likely to be inclined to use them.**[1]

This view is strengthened by the findings of Riis et al[3], which I discuss in more detail below. In his 2014 paper, Rakic notes that this could be considered a genuine concern when considering the possible viability of a voluntary programme of moral enhancement, recognising a further argument that undergoing such an intervention could even prove problematic for those that do choose to have it as their enhanced levels of empathy may lead them open to manipulation by the unenhanced. As a solution, he suggests that “external stimuli” in the form of incentives should be offered in order to encourage people to undergo the moral enhancements^{iv} and to lessen the risk of experiencing more difficulties in life as a result. Suggesting the state as a provider, he proposes the following incentives: “tax reductions, schooling allowances for their children, retirement benefits and affirmative action policies that favour them”[2].

However, the use of incentives in the context of moral enhancement raises concerns about coercion and justice, and indeed questions as to whether people would be inclined to enhance those traits likely to be affected by the endeavour to begin with. I will consider each of these matters in turn.

Coercion and Justice

One concern that could be raised relates to the impact that incentivising moral enhancement could have on voluntariness: Could the voluntary aspect be lost with the inclusion of incentives? In short, could the offer of financial incentives^v prove coercive?

While there is currently no programme in place that is directly analogous to the incentivisation of enhancement interventions, programmes of incentivising participation in research[6], adherence to drug regimens[7], and implementation of healthier lifestyles[8] could provide some insight into questions surrounding the ethics of using financial incentives in such a manner.^{vi} However, while they are similar enough to hypothetical incentivised moral enhancement programmes to provide a reasonable point of reference for this discussion due to their potential to contribute to public good (especially in the case of the first of the three), the inevitable differences between these should not be entirely dismissed. For instance, participation in research could involve greater risk to one’s

ⁱⁱⁱ Some possible reasons for this may be that people may not realise that they are morally deficient, or it may even be the case that they do realise this but are either indifferent to the fact, or even enjoy this aspect of their personality and so do not wish to change (this could be the case with career criminals, for example, or perhaps even some powerful businessmen (see[4])).

^{iv} An idea also suggested by Tonkens in[5]

^v While not all of the incentives suggested by Rakic are financial in nature, due to space constraints this paper will focus on financial incentives in particular.

^{vi} It would be prudent to note that in this paper I will be focusing my attention only on the question as to whether financial incentives could prove coercive for free persons (that is, those that are not incarcerated) in particular. There are of course various other demographics that would be affected in different ways by such proposals, but I do not have the space to discuss this in the depth that such a discussion would warrant.

health than taking enhancers that have already been regulated for public use. The implementation of a healthier lifestyle could be done without reliance on interventions such as pills or injections. So in the absence of an alternative programme that would be directly analogous with incentivised moral enhancement, these are the best possible reference points for this discussion, but one should keep in mind these relevant differences throughout the discussion that follows.

Coercive threats would present a clearer case of coercion as it is typically understood: wherein options are removed, as this is an obvious example of a coercive tactic[9]. However, incentives would seem to present an offer rather than a threat. As McMillan notes:

Threats attempt to remove options by making at least one of them undesirable and therefore sit naturally alongside the idea of coercion, which also implies that choices are rendered involuntary. Offers, on the other hand, tend to create options that otherwise would not exist.[9]

Nevertheless there is some controversy as to whether offers can indeed be considered to be coercive under certain circumstances. Shaw, for example, certainly believed that some offers could constitute coercion; when responding to a proposal from NICE which recommended financial incentives to promote adherence to methadone drug regimens[10], she expressed concern that the proposal was coercive in nature. As she put it: “by carrot rather than the stick, but coercion none the less.”[7] This view is seconded by Wiseman[11] and Selgelid[12]. Referencing Rakic’s suggestion directly, Selgelid notes:

Rakic suggests the possibility of incentivising moral enhancement, but he fails to acknowledge that this might detract from freedom to refrain from taking incentivised action – just as the threat of being told you will be shot if you don’t do something detracts from the freedom not to do it. The greater the costs of not doing something, the less free we are to do otherwise. Forgone rewards count as costs.[12]

Wertheimer and Miller, however, disagree: “The claim that the offer of financial payments can actually constitute a coercive offer in a manner that undermines informed consent is both false and incoherent, because *genuine offers cannot coerce*.”[13] They do assert that sometimes an offer can affect the decision-making process of an agent, but that this does not make the offer coercive, rather this may make it an instance of undue inducement. That an inducement could be considered undue is not an inherent feature of that inducement itself, rather whether it is considered as such will depend on the agent’s response to that inducement. As Wertheimer and Miller note:

An offer is not problematic if it is *genuinely* too good to refuse. It is problematic if it *seems* to be too good to be refused and would be refused if the agent’s judgement were not blinded or clouded or impaired.[13]

This idea is echoed by Tishler and Bartholomae^{vii}, who assert that whether a financial incentive is considered undue will vary from person to person depending on the value that each person attaches to money and the strength that they perceive a financial incentive to have. Therefore, it is unclear that a financial incentive could be objectively considered to exert undue influence.

^{vii} As cited in [14]

With this in mind, it seems reasonable to raise concerns of social justice: that people from socioeconomically disadvantaged backgrounds could be considered more at risk of being affected by undue inducements. As Permuth-Wey and Borenstein note: “financial remuneration may be more attractive to economically disadvantaged populations.”[14] Some empirical research however suggests that concerns of this nature may be misplaced; Halpern et al[6] found that with regard to willingness to participate in clinical trials – while neither the poorer nor the wealthier group were immune from the lure of financial incentives – wealthier people were far more likely to be enticed by increased levels of pay (and therefore more willing to participate) than those from a more economically disadvantaged background^{viii}.

However, it is not clear whether these findings represent a state of affairs throughout society as a whole^{ix}, nor does it seem to be the most commonly taken view. For example, Macklin writes of a medical school in an area with a large number of “urban poor [which] has a policy of not advertising or attempting to recruit subjects [for research] from the poor section of town.”[15] Further, that there is “a general suspicion of efforts to solicit subjects from among community residents because of the fear that monetary payments will serve as an undue inducement.”[15]

While their findings are controversial, the research by Halpern et al nonetheless indicates further that the question as to whether using incentives to encourage moral enhancement could be considered a form of coercion, or an undue inducement, is not one that can be straightforwardly or objectively answered. However, that does not mean that the matter can be dismissed entirely – these concerns indicate that caution is necessary when constructing policies involving financial incentives.

If the possible implications of incentivisation in relation to coercion and justice are unclear, other possibilities raise more important reasons to doubt the benefits of incentivisation. Thus, I turn now to the more pressing concern that could be raised against incentivised accounts of moral enhancement: taboo trade-offs and fundamental traits.

Taboo Trade-Offs and Fundamental Traits

My main argument against incentivising programmes of moral enhancement comes from the phenomenon of taboo trade-offs (a concept that I shall define in more detail shortly) and the findings of Riis et al that people appear unwilling to enhance those traits that they perceive to be fundamental to the self. Taken together, it seems that such a proposal for incentivised voluntary moral enhancement would be met with public indifference at best and public outcry at worst – as I shall now explain.

Parke et al noted in a study regarding public attitudes to policies incentivising healthy behaviour: “incentives may represent an attempt to put a price on something that many feel ought to be priceless.”[8] But what would this ‘priceless’ something be in the case of voluntary moral enhancement? An answer to this could be provided by Riis et al and their studies on personal

^{viii} This was demonstrated by willingness-to-participate statistics of 37% (rich) vs. 20% (poor) after an increase in the financial incentives offered.

^{ix} More research would likely be required to ascertain this.

identity and enhancement, and in particular their findings relating to the attitudes of people regarding those traits perhaps most likely^x to be affected by moral enhancement: empathy and kindness. I shall come back to this in more detail shortly, but first I shall briefly explain what is meant by the term ‘taboo trade-off’.

McGraw, Schwartz, and Tetlock define taboo trade-offs as being those that “entail comparisons of the relative importance of secular values (e.g. money, time, and convenience) with sacred values that are supposed to be infinitely significant.”[17] Taboo trade-offs are often spoken of with reference to Fiske’s theory of the four relational modes: the four types of relationships that help people to navigate and organise most social interactions and even attitudes. The four types are: Communal Sharing (people put in what they can and take as they need - found in communities of any size), Equality Matching (usually found in friendships, where tit-for-tat reciprocity of favours is commonplace), Authority Ranking (e.g. army ranking systems), and Market Pricing (that which underlies capitalism and essentially enables us to put a price on those things on which we put a price)^{xi}.

In this context, taboo trade-offs are those which occur when norms from one model are brought into another, especially when those norms are of the Market Pricing domain^{xii}. As Shiell et al note:

If a comparison across normative boundaries is attempted, as might be the case when one tries to assign a monetary value to friendship or loyalty or health for example, then one undermines the very thing that one is attempting to value. In this view, one cannot be a true friend if one is willing to value one’s friendship in monetary terms.[19]

McGraw et al[17] note that research indicates that typical responses to taboo trade-offs tend to be rather strong, with common reactions including moral cleansing and moral outrage.

It is not clear that the relationship models would be entirely helpful in the context of incentivised voluntary moral enhancement, but the principle in itself could be clearly seen to be applicable. As already noted above, Parke et al adopted the term for a similar purpose when they stated that incentives can at times appear to be an attempt to put a price on something that could be considered priceless. As I have already mentioned, the findings of Riis et al suggest that the very traits likely to be affected by moral enhancement may be considered fundamental to the self – and therefore priceless. I will give a brief account of these findings before explaining their connection to the phenomenon of taboo trade-offs, and then finally considering possible counterarguments that could be raised against this position.

Research by Riis et al demonstrated “that healthy young people are more reluctant to enhance traits that are believed to be fundamental aspects of their self-identities than traits that are believed to be less fundamental.”[3] Riis and colleagues constructed a study wherein participants were given a list of 19 traits and were asked to rate how relevant each one was to self-identity; participants were

^x Especially if we take the view of Persson & Savulescu (as asserted in a number of places, but especially in[16])

^{xi} A detailed account of these relational models can be found in[18]

^{xii} This is evidenced throughout the studies detailed in[17]

then asked to indicate whether they would be willing to enhance each trait. The results of two of the traits listed in the study are of particular significance for our purposes: Empathy and kindness. These were regarded as being the most fundamental to self-identity out of the 19 traits listed^{xiii} with only 13% and 9% (respectively) of the participants were willing to enhance these traits – the lowest figures for the entire study.

For this reason it is possible that Rakic's proposal of incentivised moral enhancement presents an example of a taboo trade-off. For if it is indeed the case that people consider traits such as empathy and kindness – which of course would likely be affected with the use of moral enhancement – to be fundamental to their personal identity, to such an extent that it makes them unwilling to enhance those traits, then it would not be an illogical assumption that people would be inclined to consider these traits priceless. Therefore, an offer of financial incentives to encourage people to undergo moral enhancement – of those traits that Riis et al's research suggests that people consider most fundamental – could be a clear case of "an attempt to put a price on something that many feel ought to be priceless"; a taboo trade-off.

One counterpoint could be raised at this juncture: that those participants (and indeed those who share this view) who believe there to be fundamental, unchanging aspects of the self are mistaken^{xiv}. However, this counterpoint does not stand in this context. For whether these people are right or wrong in their assumption that there are traits that are fundamental to the self the fact is that they do have this assumption and it is going to influence their decision when considering voluntary moral enhancement.^{xv}

A further counterpoint could be taken from another study by Riis et al reported in the same paper as the findings above. The study in question suggests that people seemed to consider enhancing traits that they regard as fundamental to their personal identity to be far more acceptable if the interventions were reframed and marketed as being an 'enablement' rather than an enhancement (perhaps appealing to a notion of unlocking potential). Illustrating this with a real-life example, Riis et al note:

In this light, it is interesting to consider the case of Paxil, an antidepressant sold by GlaxoSmithKline. Paxil has used the tagline "Paxil gets you back to being you" on its Web site. This tagline can, appropriately, ease the concerns of clinically depressed and anxious individuals who are considering taking this potentially helpful medication. At the same time, our research suggests that it could also increase the inclination of non-clinical individuals to seek a Paxil prescription for self-improvement purposes.[3]

With this in mind, one could argue that Riis and colleagues' findings regarding unwillingness to enhance fundamental traits need not prove too problematic for incentivised programmes of moral enhancement – so long as those programmes were reframed as 'enablements' rather than enhancements. However, while such a marketing strategy may allay concerns of losing one's

^{xiii} 1.38 and 1.39 respectively

^{xiv} One view that would take such a position is offered by Glover in[20]

^{xv} It could be argued that counselling prior to enhancement would allow peoples' fears on this matter to be allayed, however this assumes that people would be willing to engage with the matter in that depth – unlikely given that, as noted above, reactions to taboo trade-offs typically involve moral outrage[17].

identity, there are two reasons to doubt whether such tactics would prove sufficient in the case of moral enhancement.

First of all, it is not entirely clear that such a marketing approach would work in the context of moral enhancement. The ‘fundamental’ trait that was used in the study by Riis et al was that of Social Comfort (defined as the “Tendency to feel comfortable when meeting new people”[3]) – a trait which people had deemed to be rather less fundamental to the self than empathy or kindness^{xvi} and, correspondingly, which people were far more willing to enhance than those that we would associate with moral enhancement^{xvii}.

Secondly, even if this marketing approach (of reframing enhancement as ‘enablement’) could work in this context and was implemented, given the financial incentives on offer it could perhaps be a concern that people might become suspicious – *if it’s so good for me, why are they offering me so much? What’s the catch?*

Studies indicate that people are aware that higher incentives indicate higher risk in clinical trials[21]. It is not a great leap to assume that these suspicions and this kind of thinking could be relevant and applicable to incentivised voluntary moral enhancement, especially when it is being pitched as something that is a benefit to the enhanced (or rather, ‘enabled’) person specifically.

It could be argued that if the incentives themselves were to be removed from an account of voluntary moral enhancement, and instead the focus was shifted to employing the ‘enablement’ angle in advertising the intervention, that this would side-step the issues noted above. To merely offer a different marketing strategy such as the ‘enablement’ approach^{xviii} and remove the incentivisation aspect would simply leave us with a better-promoted version of ordinary voluntary moral enhancement. A prospect that seems to be far more uncontroversial than that of incentivised voluntary moral enhancement, which poses a significant cause for concern.

Conclusion

Given concerns that people would not otherwise be inclined to undergo moral enhancement voluntarily, incentives do at first appear to be a necessary addition to any proposed programme of the endeavour. However, while they do not necessarily result in coercion or undue inducement (which one may typically associate with the use of incentives in general), the use of incentives in the context of moral enhancement raises unique concerns, possibly leading to moral outrage in their attempt to entice people to alter characteristics that many consider fundamental to their identities.

Instead, a better approach to address the possible lack of participation would perhaps be to invest in improved marketing strategies (such as the ‘enablement’ reframing tactic mentioned above) and reject incentives for moral enhancement altogether.

^{xvi} Social comfort had a ‘identity index’ score of 0.71 as opposed to 1.38 for empathy and 1.39 for kindness

^{xvii} 26% as opposed to 13% for empathy and 9% for kindness

^{xviii} Assuming that the enablement approach could be made to work in this context, and further that no advertising ethics (e.g. exaggerating to the point of lying – therefore threatening informed consent) would be breached in this.

Acknowledgements

I am extremely grateful to Arts and Humanities Research Council and the University of Manchester for their support. I would also like to express my gratitude to Ruth Chadwick and Alexandra Mullock for their comments, advice, and encouragement in preparing this paper.

References

1. Persson I, Savulescu J. The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity. *J Appl Philos* 2008;25(3):162-77.
2. Rakic V. Voluntary moral enhancement and the survival-at-any-cost bias. *J Med Ethics* 2014;40(4):246-50
3. Riis J, Simmons JP, Goodwin GP. Preference for Enhancement Pharmaceuticals: The Reluctance to Enhance Fundamental Traits. *J Consum Res* 2008;35:495-508
4. Wasserman D. When bad people do good things: will moral enhancement make the world a better place? *J Med Ethics* 2014;40(6):374-5
5. Tonkens R. 'My child will never initiate Ultimate Harm': an argument against moral enhancement. *J Med Ethics* 2015;41(3):245-251
6. Halpern SD, Karlawish JHT, Casarett D, Berlin JA, Asch DA. Empirical assessment of whether moderate payments are undue or unjust inducements for participation in clinical trials. *Arch Intern Med* 2004;164:801-3
7. Shaw, J. Is it acceptable for people to be paid to adhere to medication? No. *BMJ* 2007;335:233
8. Parke H, Ashcroft R, Brown R, Marteau TM. Financial incentives to encourage healthy behaviour: an analysis of UK media coverage; *Health Expect* 2011;16(3):292-304
9. McMillan JP. The kindest cut? Surgical castration, sex offenders and coercive offers. *J Med Ethics* 2014;40:583-90
10. Kendall T. Paying patients with psychosis to improve adherence, *BMJ*. Published Online First: 22 October 2013. doi: 10.1136/bmj.f5782
11. Wiseman H. SSRIs and Moral Enhancement: Looking Deeper, *AJOB Neuroscience* 2014;5(4):W1-W7
12. Selgelid MJ. Freedom and moral enhancement. *J Med Ethics* 2014;40(4):215-216.
13. Wertheimer A, Miller F. Payment for research participation: a coercive offer? *J Med Ethics* 2008;34:389-92
14. Permut-Wey J, Borenstein, AR. Financial remuneration for clinical and behavioural research participation: ethical and practical considerations. *Ann Epidemiol* 2009;19(4):280-285

Sarah Carter

15. Macklin R. 'Due' and 'Undue' Inducements: On Passing Money to Research Subjects; *IRB: Ethics and Human Research* 1981;3(5):1-6
16. Persson I, Savulescu J. *Unfit for the Future: The Need for Moral Enhancement*. Oxford: Oxford University Press 2012.
17. McGraw AP, Schwartz JA, Tetlock, PE. Taboo Trade-Offs, Relational Framing, and the Acceptability of Exchanges. *J ConsumPsychol* 2005;15(1):2-15
18. Fiske AP, Tetlock PE. Taboo Trade-offs: Reactions to Transactions that Transgress the Spheres of Justice', *PolitPsychol* 1997;18(2):255-97
19. Shiell A, Sperber D, Porat C. Do taboo trade-offs explain the difficulty in valuing health and social interventions? *J Socio Econ* 2009;38:935-939
20. Glover J. I. *Philosophy and Psychology of Personal Identity*. London: Penguin Books Ltd 1991
21. Cryder CE, London AJ, Volpp KG, Loewenstein, G. Informative inducement: Study payment as a signal of risk. *Social Science and Medicine* 2010;70:455-64