

See discussions, stats, and author profiles for this publication at: <http://www.researchgate.net/publication/281901398>

Moral Enhancement: Do Means Matter Morally, Reply to Justin Caouette

ARTICLE · AUGUST 2015

READ

1

2 AUTHORS:



Farah Focquaert

Ghent University

28 PUBLICATIONS 109 CITATIONS

SEE PROFILE



Maartje Schermer

Erasmus MC

77 PUBLICATIONS 539 CITATIONS

SEE PROFILE

Online Neuroethics Symposium on The Brains Blog

<http://philosophyofbrains.com/2015/08/25/neuroethics-symposium-on-focquaert-schermer-moral-enhancement-do-means-matter-morally.aspx>

Moral Enhancement: Do Means Matter Morally?

Reply to Justin Caouette

Farah Focquaert & Maartje Schermer

We would like to thank Justin Caouette for his interesting comments and for given us the opportunity to further clarify our views. In his comments, Caouette addresses three points. First, he questions whether and why reflection and deliberation are necessary for authenticity of behavior. Caouette argues that we do not deliberate about everyday psychological changes (e.g. in our mood) either and that this is not a problem for the authenticity of our behavior, as long as we can still relate to such changes in some way, and can choose how to respond to them, and how to act. Our concern in the paper, however, is more with authenticity of personality and identity, than with isolated ‘behavior’. Actually, the point Caouette makes here with regard to behavior - that as long as we have the ability to ‘step back’ and reflect on ourselves, on our motivations, and choose to endorse or reject them, there is no problem of authenticity – comes close to the point that we have argued for in relation to identity.

We agree that as long as we talk about relatively minor changes like ‘normal’ fluctuations in mood, for example, and a person remains, by and large, ‘himself’, with the same values, goals, desires and traits, he can indeed reflect on these changes, and choose how to respond to them, or how to act. However, we have argued that some passive, direct intervention may cause more radical changes, and these might affect the person to such an extent that he is, in a sense, not the same person anymore. If the intervention changes his values, desires, propensities or his outlook on life radically, then his authenticity as a person, the authenticity of his personal identity, is compromised. Caouette writes that even if certain ‘psychological states’ are implanted in us, ‘we’ remain the source of what we decide to do with our mental states. However, if ‘we’ ourselves have changed considerably, it is not clear anymore that our choices will really be authentically ours¹. Moreover, we have argued that some passive interventions effectuate changes that a person is *not aware of*, and hence, cannot really reflect on or take a stance towards. If a person does not really notice he is in a lousy mood, he will not be able to choose how to deal with that or respond to it. That is why we have argued that *concealed* narrative identity changes are particularly problematic.

Further on in his comment, when discussing identity, Caouette questions whether, and why, reflection and endorsement regarding the genesis of one’s desires, values and personality traits is

¹ By the way, the idea that we will always have the power to just decide what we will do with our mental states, seems a simplification as well. People who are addicted, seriously depressed, manic etc. do not have the power to simply ‘choose’ what to do.

necessary for narrative identity. Here there appears to be another misunderstanding, since we do not claim it is necessary for having some sense of narrative identity, or for being able to tell a coherent story about one's life, but for having an authentic narrative identity.

Here it may help to explain that our conception of authenticity, and the one used by DeGrazia, is in line with a theoretical tradition of authors such as Frankfurt, Dworkin and Christman. In this line of thought a person is autonomous or authentic² when he leads his life in accordance with his own goals and values, and his actions are authentic when he acts on preferences that are in line with his second order volitions, that are in turn consistent with his overall values and goals. Whether a person's goals and values are truly his in the relevant sense, depends on whether he has endorsed them after rational reflection. Only endorsing certain values, however, is not by itself enough to make them truly 'one's own'. If it were, all forms of brainwashing, manipulation, subliminal suggestion and the like, that succeeded in making the person in question 'endorse' certain values and desires, would count as authentic. Dworkin called this 'illegitimate influences' and claimed that identification (or endorsement) should not itself be influenced in ways which make the process of identification alien to the individual. So, it matters how one's endorsement came about: through the application of one's own moral reasoning, or through external manipulation.

We would like to focus the remainder of our reply on the third comment raised by Caouette. Given the fact that we already restrict the autonomy of individuals in prison and one of the goals of incarceration is to alter one's moral compass, Justin Caouette argues that the concerns we voice with respect to forced (forced) direct interventions are overblown or even moot because 'we' do not value the identity of those individuals who identify strongly with certain counter-moral personality traits. One of the main goals of writing our article was to prevent unethical applications of passive interventions such as forced administration of direct neuro-interventions (for moral enhancement) within the criminal justice system. There are a lot of criticisms that come to mind when we read this section of his commentary, and we cannot discuss all of them within this space. We will focus on three main issues.

First, who is the 'we' in the argument put forward by Caouette? Is it society or the criminal justice system or the philosophical community who does not value the identity of offenders with certain counter-moral identity traits? It is unclear to whom the 'we' refers and if this 'we' has any relevant authority or expertise with regard to punishment. Offenders are persons with basic human rights which need to be respected similar to the basic human rights of law-abiding citizens. For example, offenders have a basic human right to receive adequate health care, including mental health care, and a right to mental integrity. Offenders, even if the crimes they have done are horrific, should be treated with respect for their autonomy and identity insofar as the protection of society allows this. No harsher measures need to be taken than those that are sufficient to adequately protect society (Pereboom, 2014). Of course, in case of serious crimes or repeat offending, society has a duty to protect its citizens and adequate safety measures need to be taken, such as (potentially indefinite) confinement. However, forced (direct) interventions do not respect human dignity and autonomy, and may conflict with the 'do no harm' principle. We have

² Unfortunately these terms are not used in the same way by different authors.

not argued that either direct or passive interventions are never justified, but they should be voluntary (i.e. with informed consent). Farah Focquaert and Sigrid Sterckx have argued elsewhere which other conditions need to be met in order to render offering (direct) neuro-interventions to offenders ethically permissible (Focquaert, 2014; Focquaert & Sterckx, in press). Forced (direct and/or passive) neuro-interventions are generally ethically problematic.

Autonomy can be defined in a variety of ways. Autonomy can focus on negative freedom, positive freedom, authenticity, empowerment, etc. It is true that we restrict offenders' autonomy when we incarcerate individuals in the sense of taking away offenders individual liberty, which is a fundamental human right. In order to sanction offenders by incapacitating them, by taking away their individual liberty, governments need to make sure that no other sanctions would work equally well in addressing their goals. In general, due to today's over-crowded prisons and the lack of available resources to provide adequate care, community-based sanctions are much less conducive to human rights violations than imprisonment and should therefore be the preferred option for most offenders provided that the rights of victims are upheld and the safety and concerns of society at large addressed (Focquaert & Raine, 2012).

Second, reducing an individual to certain aspects of his/her being that are considered socially undesirable is a form of stigmatization. We fully agree that the moral wrongness of deviant behavior needs to be identified, but this need not be done by reducing individuals to certain aspects of their behavior and thereby labelling and stigmatizing individuals (e.g. as sexually violent predators). Such labeling and stigmatization also occurs when someone with obesity or an alcohol addiction gets reduced to his/her (medical) condition and gets portrayed in a negative way, or when signalling out individuals with serious infectious diseases as those who should be avoided at all cost.

Only 15 to 25 percent of prisoners have psychopathy, and these individuals indeed suffer from a serious personality disorder exemplified by counter-moral traits. However, most offenders do not suffer from such a disorder and do not commit serious crimes such as murder, rape and physical assault, or massive fraud. It is one thing to say that we do not value counter-moral personality traits and/or behavior, but quite something else that we do not value the identity of individuals with certain counter-moral personality traits. If Caouette merely focuses on those offenders who strongly identify with such traits, his claim is limited to a small group of offenders of which some will identify with such traits continuously while others will be willing and able to change their future behavior. Most offenders however do not experience or identify their crimes as constitutive elements of their identity, nor can we regard this to be the case from a psychiatric or psychological perspective.

Third, the argument on forced interventions that Caouette voices is set against the alternative of (mass) incarceration. It is our view that mass incarceration and the conditions under which many offenders are held today are not ethically warranted in themselves. For example, a major problem that prisoners face upon re-entry is that their time spent in prison has eroded their life both professionally (lack of training, loss of skills, limited or no employment opportunities), and privately (family and friends, emotional support networks in general, housing opportunities). Moreover, stigmatization acts as an indirect, enduring form of punishment that further aggravates

the situation. Not surprisingly then, recidivism rates are extremely high. In fact, imprisonment is found to be counterproductive in terms of rehabilitation and reintegration in society for low-level offenders and for those in need of special care, such as individuals with mental health and substance abuse problems. Moreover, the long-term costs of imprisonment on offenders' family members and communities are often neglected (Focquaert & Raine, 2012). Hence, we feel that neither mass incarceration nor forced interventions are the most effective and/or ethically acceptable means to address deviant behavior. The fact that incarceration may harm convicts – as Caouette agrees they do – is no reason to opt for other harmful interventions such as forced direct intervention, but to look for alternatives.

A brief final note: like Elizabeth Shaw, Justin Caouette seems to have understood us as saying that direct interventions are always passive, and indirect interventions are active. As explained in our response to Shaw, this is not what we have claimed – but perhaps we have not been clear enough on that point, so we thank the commentators for giving us the opportunity to clarify ourselves!

References

Focquaert, F. (2014). Mandatory neurotechnological treatment: Ethical issues. *Theoretical Medicine and Bioethics*, 35, 59-72.

Focquaert, F., & Raine, A. (2012). Ethics of community-based sanctions. S. Barton-Bellessa (Gen. Ed.), *Encyclopedia of community corrections* (pp. 144-148). SAGE Publications.

Focquaert, F., & Sterckx, S. (in press). Offering neurointerventions to offenders with impaired moral functioning: Ethical and criminal justice aspects. To appear in Nicole Vincent (Ed.), *Neurointerventions and the Law: Regulating Human Mental Capacity*.

Pereboom, D. (2014). *Free will, agency and meaning in life*. Oxford University Press.

On the Moral Permissibility of Passive Moral Enhancement*
Comment on “Do Means Matter Morally?”
Justin Caouette
University of Calgary

In this short response to Farah Focquaert and Maartje Schermer’s “Moral Enhancement: Do Means Matter Morally?” I’d like to focus on two general issues discussed in the paper regarding the use of direct, passive interventions to morally enhance subjects that have impairments in moral decision making and/or possess counter-moral personality traits. First, I’d like to consider the perceived threats to authenticity raised by the authors. I’ll argue that the threat is not apparent given that these enhancements, presently hypothetical as pointed out by the authors, do not determine their subjects to act in any particular way. Given that the subjects will still have free will, it is suspect to assume that the behaviour stemming from decisions after enhancement are not authentic. Next, I’ll briefly consider the threats that the authors raise regarding “concealed narrative identity changes”. I’ll suggest that these concerns might be overblown, though they are important to consider when thinking about the permissibility of direct, passive interventions. Lastly, I’d like to focus on the enhancement of criminal offenders. Given that we already restrict the autonomy of those in prison and given the goal of incarceration is to alter one’s moral compass, I will suggest that the concerns raised by the authors, even if not overblown, might be moot given that we do not value the identity of those who identify strongly with the counter-moral personality traits that these offenders exhibit.

1. *Threats to Authenticity*

Focquaert and Schermer argue that there is a morally relevant difference between what they call active and passive interventions. On their view, active interventions seem less likely to threaten identity and autonomy because of the nature of the intervention. Active, indirect interventions entail that the agent is involved in the process of changing their brain, which usually takes the form of therapy or moral education. This process is usually much longer and requires that the agent undergoing the change has room for rational reflection and deliberation about the changes that are under way, as well as an active engagement from the person undergoing the change. On the other hand, the authors indicate that direct, passive interventions create a greater cause for concern. Passive interventions (i.e DBS, TMS, and tDCS) have a greater potential to compromise identity and raise intrinsic doubts concerning authenticity and autonomy.³ It is claimed that this is so because without the active engagement from the person undergoing the DBI we bypass conscious reflection, continuous rational deliberation, and autonomous choice. This reflection, deliberation, and choice is assumed by the authors to be essential for the behaviour to be authentic to the enhanced subject. This assumption is what I’d like to hear more

³ Nicole Vincent raises similar concerns in her article “Restoring Responsibility: promoting justice, therapy and reform through direct brain interventions”, *Criminal Law and Philosophy* (2014) 8: pages 21-42.

about. Why think that directly altering an individual's brain functioning bypass one's authenticity? Let me offer some reasons to believe that the threat to authenticity might be overblown by the authors.

First, is it the case that we ever deliberate about the psychological changes we find ourselves with from day to day? For instance, consider waking up in the morning and finding yourself edgy. We might reflect on why it is we are edgy, but after deliberating we might not find the source. This seems plausible. But just because we find ourselves edgy it doesn't follow that we are forced to let that edginess win the day, so to speak. We can still deliberate and choose to do things that alter our mental state. We may try watching a comedy or listening to music that helps us overcome our edginess. We may decide to take a drug to remove the edge. These choices are available to us even though we have no idea why we are edgy. The point of bringing this to light is to show that a changed psychological state, even a state that we do not know the source of, does not determine us to act in any particular way. We are still the source of what we decide to do with our mental states. Even though finding ourselves edgy out of the blue will have an affect as to what actions will seem salient to us, it doesn't cut off all of our options. And, given that we normally don't choose what state we find ourselves in, the authenticity of our selves and the actions that flow out of us seems to be in tact. Similarly, if the cause of our edginess is the result of some direct, passive intervention, it doesn't follow that what we decide to do about this is somehow inauthentic. External factors often put us in a variety of psychological states that we do not choose to be in. For these reasons I would like to hear more from the authors regarding why it is that direct, passive interventions pose any further threats to our authenticity than the external factors that often put us in a variety of psychological states.

I believe authenticity is important when thinking about our actions. If an action is authentic, we must be the source of that action - it must be up-to-us, so to speak. But to use the language of authenticity with regards to our psychological states seems to raise a bunch of issues. If one admits that our psychological states are often altered and produced, in part, by external factors, then I think we need to hear more as to why direct, passive interventions pose greater risks to the authenticity in question. It seems that our actions can still be authentically ours even if some of the psychological states or newly implanted ethical alternatives are partly the cause of our action. After all, external factors often put us in a variety of psychological states but we do not think that the actions that arise from these states are inauthentic so why think that the actions that arise from a person that has been morally enhanced, given that many of these factors we are unaware of, are inauthentic?

2. *Threats to Identity*

The authors claim that "to be sufficiently aware of one's narrative identity, a narrator must be able to coherently 'explain' or communicate the narrative unfolding of her life story". I don't want to challenge this claim. Instead, I'd like to raise some questions as to why it is that an individual enhanced directly via passive interventions cannot coherently explain or communicate the narrative unfolding of her life story. Let's consider the case of Manuel.

Manuel never acts morally; he is a person who doesn't consider the feelings of others and takes pride when others point out to him that he is not a person who cares about the well-being of others. One night while sleeping, Manuel's mom morally enhances him directly and passively in a way that makes him more empathetic and he is unaware of this enhancement. Manuel is aware of the change though: when he reflects on the person he is, much to his surprise, he finds himself caring about others more. Why does he lose the ability to speak to this change? And if he doesn't, then why think his narrative identity is threatened? Is it typically the case that we know how and why our feelings change, or why we mature or have an epiphany?

It is true that we may be able to tell a story as to why we change sometimes, pointing to some specific experience that gave rise to the change in feeling or disposition. However, this is often not the case. One might wake up by the side of their spouse whom they loved and wanted to be with all of the days prior to this one, but on this day could realize they do not want to be with them anymore. It happens quite often and sometimes occurs without the person changed knowing why. We can try to make sense of why we suddenly feel this way but it's not clear to me that we can always know why. More importantly, this lack of knowledge doesn't seem to threaten our narrative identity. As long as we have the ability to endorse or choose not to endorse the way we feel at a certain time it seems that our narrative identity is in tact.

In sum, the authors endorse DeGrazia's theory of personal narrative which they describe as depending on two things: "whether the person identifies with her new traits or desires; and whether she identifies with their genesis, i.e., the way they came about." But I'd like to hear more as to why this second condition must be met. In other words, why should we buy DeGrazia's theory? Why isn't enough to meet the first condition to say that we have a narrative identity?

I believe abrupt changes can be incorporated into one's narrative. And further, that identity should be understood as a cohesive narrative. One can identify with their new self as a direct result of the DBI that was forced on them due to choices *they made*. Further, if one believes that the self is dynamic and always changing with new experiences this would not be a big bullet to bite. This would require that they have knowledge of the intervention; however, it doesn't follow that they must have consented to it in order for it to be incorporated into one's narrative. It is a part of my narrative identity that my two parents raised me and that I was punched by a bully in middle school. But I didn't consent to either. Similarly, one who has been morally enhanced might be more moral (which would be a distinct difference to the person who didn't care about morality prior to the enhancement) and might understand that this is a result of the state forced DBI imposed on them from their own transgressions. This is not problematic because they would still have memories of their past self and those memories, and coupled with their new take on moral situations, this would give rise to a distinct self. So, while these new salient moral decisions might seem alien, they could easily identify with them and endorse them once they recognize that their new found moral awareness will open doors human flourishing that had previously been closed. Further, such drastic changes to one's self is no different (morally speaking) from many transformative experiences that we often go through without consent (I'd point to the work of Laurie Paul as evidence for these sort of situations).

3. *Direct, Passive Enhancement on Criminals*

Before concluding I would like to turn to a specific case of moral enhancement, the enhancement of convicted criminals. For example, let's consider DBI's that aim to morally enhance persons convicted of a crime. The authors present concerns related to a person's autonomy, authenticity, and identity, which in turn raises serious doubts about the moral permissibility of such interventions in general. But, given that many who will need these enhancements are likely to be the criminal population, I wonder if their concerns apply to a population that already has limited autonomy and whose identities we do not respect.⁴ Here I'm thinking of the person who identifies as a rapist or a person that identifies as the type of person who will assault another for looking at them "funny". Do we respect such identities? Do we care if folks cannot recognize themselves as **those** types of people anymore if they were morally enhanced? It's doubtful. Even if we admit that a convict's self is abruptly and drastically changed this need not trouble us. Implemented properly, DBI's can bring about benefits for society at large and/or for the convict themselves. Next, even if we agree that abrupt changes to the convict might be alter their identity, this abruptness of change need not be considered negative, especially for those individuals who identify with harmful and criminal behaviours. Why value the identity of someone who identifies as a racist or as someone who beats women (assuming they have been convicted of crimes relating to these world views)? We already restrict the scope of a convict's autonomy when sentencing them to prison or mandated rehabilitation. We do this because we deem that identifying as such is not a worthy identification. And further, if someone identifies with morally bankrupt traits we try our best to eradicate them, short of methods that could harm the criminal. I would be interested to hear from the authors whether making someone morally better can be considered harming them.

Further, we could ask why think that the direct, passive method is worse? Comparing traditional rehabilitative methods or active, indirect interventions, it seems an argument can be made that both are just as likely to cause problems to identity in criminal offenders. In fact, there might be more reasons to be concerned with the longer process when thinking about identity and autonomy. Change likely occurs in both active and passive interventions, however, in the case of active interventions, changes are accompanied by some negative side effects (which are less likely to occur with passive, direct interventions). Given the amount of time that convicts spend with others who are not committed to changing their behaviour, this forces those who might take an active role in their cognitive changes to build defense mechanisms that also shape who they are. Given that these mechanisms can become part of one's character, it's worth investigating how much this can negatively affect their future relationships and life plans. It doesn't seem far fetched to assume that the active, indirect interventions carry possible harms to the criminal offender that direct, passive interventions do not. This suggests that direct, passive interventions might be a better choice when considering all the pros and cons when deciding how to rehabilitate or morally enhance our convicted felons who identify as rapists, racists, and those quick and ready to assault folks who pose no physical threat to others.

⁴ Bublitz and Merkel consider this population in their 2014 paper "Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination", *Criminal Law and Philosophy* 8: 51-77.

In conclusion, not being allowed to pursue one's interests (because of incarceration and limited freedoms more generally) would seem to have a much stronger negative effect on the self than do direct, passive interventions. Having desires that cannot be fulfilled forces one to have a constant sense that their life is unfulfilled and can lead to depression and other mental disorders. Thus, having a severely restricted set of options to pursue because of forced rehabilitation or incarceration seems to be a harm that may be lessened with the use of direct, passive interventions due to the abrupt nature of these latter interventions. Given that one of the main goals of incarceration is rehabilitation, it seems that direct, passive intervention could be better at accomplishing this goal, even if we admit that one may be alienated from their prior non-moral self. Direct, passive interventions can alter one's brain so that moral choices become salient or live options. This should be seen as a positive rather than a negative effect on one's autonomy. The moral options are now live, but, they are not forced and thus their decision to endorse their new moral inclinations keeps their authenticity and personal narrative in tact. Affording one the opportunity to flourish seems better than hoping that one can come to some realization that they should see some moral action as a live option to them, especially when the track record of active traditional rehabilitation is so poor. Replacing IBI's with DBI's might allow us to extend their autonomy rather than stunt it, contra the suggestion by the authors. The abrupt change in one's cognitive apparatus might allow for them to be released much earlier than they normally would, given that the aims of the rehabilitation would have been met. I'd like to hear more from the authors regarding enhancing directly, and passively in the context of criminal rehabilitation. After all, this is the population that is most ripe for the implantation of such interventions.

*I'd like to thank David Boutland, who is my coauthor on an unfinished paper where we discuss many of these issues.