

THE EPISTEMOLOGY OF MORAL BIOENHANCEMENT

ABSTRACT: Moral bioenhancement is the potential practice of manipulating individuals' moral behaviors through biological means in order to help resolve pressing moral issues such as climate change and terrorism. This practice has obvious ethical implications, and these implications have been and continue to be discussed in the bioethics literature. What have not been discussed are the epistemological implications of moral bioenhancement. This paper details some of these implications of engaging in moral bioenhancement. The argument begins by making the distinction between moral bioenhancement that manipulates the contents of mental states (e.g., beliefs) and that which manipulates other, non-representational states (e.g., motivations). Either way, I argue, the enhanced moral psychology will fail to conform to epistemic norms, and the only way to resolve this failure and allow the moral bioenhancement to be effective in addressing the targeted moral issues is to make the moral bioenhancement covert.

INTRODUCTION

It is uncontroversial that it is both appropriate to attempt to become morally better people as well as appropriate to do so through education. The idea that we should attempt to become morally better people through biological manipulation, however, has generated controversy. This manipulation, moral

bioenhancement, is supposed by Persson and Savulescu¹ to be important in responding to the unique and potentially devastating threats that face 21st Century humans, such as climate change and terrorists whose reach and potency is increasing with advances in technology and globalism.

There is no universally accepted definition of moral bioenhancement. Raus et al.² provide a thorough and helpful taxonomy of the definitions. The definition of moral bioenhancement that I use is this:

Moral bioenhancement: The enhancement of a person's moral attitudes, motivations, or behavior through biological means.

This definition may not appear helpful, as it includes ‘enhancement.’ And there are ambiguities about what it means to enhance an attitude, typically a moral belief, or moral motivation or moral behavior.³ It is more specific than several definitions of moral bioenhancement, such as the one that defines it as the

¹ Persson, I., & Savulescu, J. 2014. *Unfit for the Future: The Need for Moral Enhancement*. Oxford. Oxford University Press.

² Raus, K., Focquaert, F., Schermer, M., Specker, J., & Sterckx, S. On Defining Moral Enhancement: A Clarificatory Taxonomy. *Neuroethics* 2014; 7(3): 263–273.

³ Agar, N. A Question About Defining Moral Bioenhancement. *J Med Ethics* 2014; 40 (6): 369–70.

improvement of person's moral capacities.⁴ A person's moral capacities are presumably exhausted by his or her moral attitudes, motivations, and behaviors. One might also wish to amend the above definition to include dispositions and emotions.⁵ Whatever one's potential protests, the above definition is general enough to capture the notion of moral bioenhancement and specific enough to be useful. Given this definition, the range of potential methods of moral bioenhancement is wide and includes the commonly expressed method of morally enhancing people by administering pharmaceuticals that induce greater empathy and altruism or virtue.⁶ There is disagreement about whether the focus of moral

⁴ DeGrazia, D. Moral Enhancement, Freedom, and What We (Should) Value in Moral Behaviour. *J Med Ethics* 2014; 40 (6): 361–368.

⁵ Jebari, K. What to Enhance: Behaviour, Emotion or Disposition? *Neuroethics* 2014; 7(3): 253–261.

⁶ Persson & Savulescu, *op. cit.* note 1; Persson, I., & Savulescu, J. Getting Moral Enhancement Right: The Desirability of Moral Bioenhancement. *Bioethics* 2013; 27(3): 124–31; Raus et al., *op. cit.* note 2; Sparrow, R. Better Living Through Chemistry? A Reply to Savulescu and Persson on “Moral Enhancement.” *J Appl Philos* 2014; 31(1): 23–32.

bioenhancement should be on individuals (e.g., just psychopaths),⁷ though I focus below on the moral bioenhancement of larger populations.

Some alleged problems with moral bioenhancement are that it is self-subverti,⁸ inequalitarian,⁹ and not currently possible.¹⁰ I don't directly engage with these problems in what follows, not because I don't think they are worth discussing, but because I suggest further difficulties that have so far gone unnoticed. I think that moral bioenhancement is mostly morally unproblematic.

Below I argue that there are epistemological difficulties with the administration of a moral bioenhancement program. These difficulties, I argue, would render a program ineffective unless the program is administered covertly. The difficulties hinge on a distinction between what is actually being enhanced. There are three aspects of person that could be manipulated. One type of

⁷ Douglas, T. Moral Enhancement via Direct Emotion Modulation: A Reply to John Harris. *Bioethics* 2013; 27(3): 160–168; 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): 162–177.

⁸ Sorensen, K. Moral Enhancement and Self-Subversion Objections. *Neuroethics* 2014; 7(3): 275–286.

⁹ Sparrow, R. Egalitarianism and Moral Bioenhancement. *Am J Bioeth* 2014 ;14(4): 20–8.

¹⁰ Crockett, M. J. Moral Bioenhancement: a Neuroscientific Perspective. *J Med Ethics* 2014; 40 (6) 370-371.

manipulation is direct manipulation of a person's behavior. A manipulation of this type is independent of a person's other states, such as his or her beliefs, dispositions, emotions, or motivations. In many cases, the manipulations of this sort would be in spite of the person's other states. I assume that a behavioral neuroscience that could manipulate a person's behavior without also manipulating another type of state is too far-fetched to be considered a legitimate candidate for the cornerstone of moral bioenhancement program.

A second type of manipulation would be a manipulation of a person's representational mental states. These are the mental states of person that represent how the world is or how it ought to be. Beliefs are the paradigmatic representational state, but other representational states are judgments, seemings, desires and all the other propositional attitudes. The mark of these types of states is that they have content and can be assessed for accuracy and, for many types like beliefs, assessed for truth. Manipulations of this type of state would manipulate the content of these states, so I call manipulations of this sort *content-oriented* manipulations.

A third type of manipulation would be a manipulation of person's content-less states, or states of a person that are non-representational.¹¹ These types of states may include affective states (though some argue that moods are

¹¹ For many, this would imply that such states are not mental states. Whether being representational is the mark of mental states is something about which the present position is neutral.

representational) and motivations. In contrast to content-oriented manipulations, manipulations of this type would be *content-free*. Depending on what one thinks about the nature of dispositions and emotions, either could be considered to be content-free or content-oriented. Dispositions are probably most naturally considered to be representational, and many people think that emotions are representational, but for the purposes of what follows nothing of significance hinges on how these types of states are best categorized.

In what follows, I assume that the manipulations of a moral bioenhancement program would be either content-oriented or content-free. More specifically, I assume that the program would manipulate either a person's moral beliefs (content-oriented) or a person's motivations (content-free). Any manipulation will eventually have to go through participants' beliefs or their motivations. For example, inducing certain emotions wouldn't contribute to moral enhancement without those emotions resulting in greater motivation to perform the prescribed actions. Putting oxytocin in the water supply to induce a greater sense of altruism won't result in a morally better society unless that sense of altruism results in participants being motivated to act on it. And if a person fails to believe that a certain action is the right thing to do, they are highly unlikely to perform it. So moral bioenhancement could proceed by manipulating lots of different things. No matter what those manipulations are, they will have to eventually go through either participants' beliefs or their motivations. The argument below, therefore, works for any other types of states that are either content-oriented or content-free.

A moral bioenhancement program could be administered three different ways. First, participation could be completely voluntary.¹² It is plausible that if moral bioenhancement were merely voluntary, so few people would volunteer that the moral enhancement of society would be insufficient to meet the the obligations required to resolve the most pressing moral issues. So I assume that an effective moral bioenhancement program would not be voluntary. Second, a program could be overt but compulsory in a way similar to how municipal water supplies are flouridated or vaccination programs administered.¹³ The administration of the program is transparent and public, and unless a person actively decides to not participate (e.g., by buying unflouridated bottled water or requesting a waiver for vaccination), participation is compulsory. Third, the administration could be completely covert and the participants totally uninformed of the program. The conclusion of the argument that follows is that because of the epistemological issues, for a moral bioenhancement program to be effective, it must be administered covertly. A program is effective only if it changes the moral behavior of a large enough number of people to meet the demands of the pressing moral issues the resolution of which is intention of the program.

CONTENT-ORIENTED MORAL BIOENHANCEMENT

¹² Rakić, V. Voluntary Moral Bioenhancement Is a Solution to Sparrow's Concerns. *Am J Bioeth* 2014; 14(4): 37–8.

¹³ Persson & Savulescu, *op. cit.* note 5.

Content-oriented moral bioenhancement is the manipulation of person's mental states with moral representational content, typically a person's moral beliefs. An example of such a manipulation may be the administration of oxytocin to a population to induce greater affection for others or the disposition to be more altruistic. This affection or altruism may result in the individuals having different moral beliefs, such as the belief that one ought ride one's bike today rather than drive to conserve fossil fuels; or the belief that one ought not eat beef this week so that fewer resources are withheld from a starving family; or the belief that one ought to give the homeless person on the way home one's pair of shoes. This may be the most currently feasible method of moral bioenhancement.

Another, perhaps less direct, method would be by administering something that makes people more sensitive to moral reasons and better at moral reasoning (even if such reasoning is not conscious).¹⁴ For example, perhaps the manipulation may cause a person to better understand and recognize patterns of human suffering, and from these patterns arrive at conclusions—moral beliefs—the contents of which are the same as those that result from the administration of oxytocin.

There may be other ways to biologically manipulate the moral content of an individual's mental states. The biological manipulation of the content of other mental states is commonplace, which, on the face of it, provides a reason to think that a content-oriented moral bioenhancement program is technologically feasible.

¹⁴ Harris, J. Moral Enhancement and Freedom. *Bioethics* 2011; 25(2): 102–111.

For example, the practice of psychiatry regularly uses pharmaceuticals to manipulate the contents of patients' mental states, such as their emotions, perceptual representations, and representations of themselves. Less direct manipulations of the content of an individual's mental states are even more common. Painkillers manipulate representations of pain (more controversially representational) and the distal beliefs about whether one is in pain, its location, and severity. Viagra manipulates (among other states) a man's belief that he wants to engage in sexual intercourse. Granted, the manipulations in these examples induce other effects, but without the manipulations the beliefs with these contents would be absent. And no belief is exempt from epistemic constraints, no matter whether it is a secondary effect.

Suppose that a moral bioenhancement program is content-oriented and overt and compulsory. The program requires that a chemical be added to the municipal water supply and this chemical manipulates citizens' moral beliefs, beliefs upon which they then act and improve the world. Because it is compulsory, unless a person goes out of his or her way to have special water filters installed or purchases bottled water without the chemical additive, his or her moral beliefs will be targeted for manipulation. And because it is overt, the citizens, or at least most of the citizens, know that this is happening; they know that their beliefs are being manipulated (in a way roughly analogous to how most people know that their water is fluoridated).

Such a program whose manipulations are content-oriented implies an epistemological difficulty that would likely render the program ineffective. The

difficulty is grounded by the fact that, generally, adult humans are aware of what they believe; our beliefs are, for the most part, transparent. Though it may be false that believing that *p* implies that one knows that one believes that *p*, or even that one believes that one believes that *p*,¹⁵ in most cases if a person believes that *p* he or she will be aware of that belief. This is just to say that he or she will believe that he or she believes that *p*.

A moral bioenhancement program that is overt and compulsory and that manipulates participants' moral beliefs can be effective only if it *changes* participants' moral beliefs. For participants whose moral beliefs are the very same as those induced by the program, the institution of the program has no bearing on those participants' moral beliefs or their subsequent behavior—they would have those beliefs anyways. So the only way for it to be effective is to change participants' moral beliefs. But given this and the fact people are generally aware of their beliefs (especially their moral beliefs), participants will be aware of the new beliefs. Combine this with the fact that they were aware of their previous, different moral beliefs, and the result is that the participants will be aware of a change in their moral beliefs. For instance, a moderately reflective participant would exhibit the following pattern of reasoning: "I used to believe that eating beef every day was okay; now I believe that it's morally wrong to eat beef; so, my moral beliefs changed."

¹⁵ Williamson, T. 2002. *Knowledge and Its Limits*. Oxford. Oxford University Press.

Now comes the epistemological difficulty. The participants will be aware of a change in their moral beliefs. But they won't have become aware of any new evidence, because, relative to the new moral beliefs, no new evidence has been introduced. By stipulation, the program changes the content of their moral beliefs. It does not change any of their reasons for that belief. Thus, participants will be aware of a change in moral beliefs but will have acquired no new evidence, which means they will lack justification for the new belief. Indeed, for the reflective participant, he or she will be aware that he or she has acquired no new evidence, the appropriate response to which is either dissent to or suspension of the belief. Participants will fall into at least one of the following two groups: those that in fact lack justification for their new moral beliefs and those that also believe that they lack justification for their new moral beliefs. Which group one falls in is a matter of one's introspective abilities. Either way, I assume, perhaps controversially, that when people lack justification to hold moral beliefs, and especially when they believe they lack such justification, they rarely act on those beliefs. If so, then when the program changes a person's moral beliefs, the change will typically fail to induce the targeted behavior. Call this the reflection problem.¹⁶

¹⁶ Another argument that the participants will lack justification for their new beliefs is that being aware of a change of belief without awareness of a change of evidence should cause one to doubt the reliability of their belief-forming faculties. And doubting the reliability of the process that produce the belief defeats the

So, either the moral bioenhancement program will not change a person's moral beliefs or it will. If it doesn't change a person's moral beliefs, then it won't be effective, because it won't matter. If it does change a person's moral beliefs, that person will lack justification to hold them and, because of this, fail to act based upon them, rendering the program ineffective. Either way, the program is ineffective.

One may object that the reflection problem requires that the participants be epistemically idealized and that this idealization doesn't reflect actual epistemic practices, and thus it is impossible to draw conclusions about what will or won't result from a moral bioenhancement program. Though I disagree that epistemic practices such as introspection and holding beliefs based upon one's evidence are idealized—they are rather common—if participants in a moral bioenhancement program fail to conform with such standard epistemic practices, then the program will have epistemic difficulties far beyond the reflection problem. One may also object to the assumption that when participants lack justification for a belief, they are unlikely to act upon it. To reject this assumption is to hold that people commonly act upon beliefs for which they have no justification. If this is indeed the case, then manipulating people to induce

justification one has to hold those beliefs. Whether the process is actually reliable is irrelevant, for the person will believe that the process is unreliable, which is sufficient to defeat justification. Thus, even if one's beliefs are produced by a reliable process, the belief that they are not is enough to defeat that justification.

particular moral behaviors will be impossible, at least if those manipulations are mediated by mental states that are supposed to conform to norms of rationality.¹⁷

¹⁷ Even those who are religious and purport to act based on religious principles have justification to hold their moral beliefs. There is a difference between having justification to hold a belief and the reasons upon which one bases his or her belief. For example, a person may read in *The New York Times* that President Obama will be in Philadelphia today. This provides one with justification to believe that Obama will be in Philadelphia today. But if the person holds the belief based upon the astrological signs, then the belief is inappropriately based, but it is still justified for the person. The fact that the basing relation fails epistemic standards does not imply that the person lacks justification to believe it. Similarly, a person who is religious and so holds beliefs that fail to conform to epistemic standards will still likely have justification to hold the moral beliefs that he or she does, as such people usually have the same sort of evidence that everyone else has (e.g., experiences of suffering, charity, etc.). It's just that they base their beliefs on the wrong evidence. So, even irrational religious moral believers will have justification to hold their moral beliefs; they just hold them for the wrong reasons. The assumption above has nothing to do with the basing relation, just what one has justification to believe. The content of the assumption is that it is usually not the case that a person acts upon a belief and that the person lacks justification to hold the belief.

The reflection problem rests on the fact that participants will be aware of a change in belief without a change in evidence. Notice that this awareness is independent of knowing or believing that one is participating in a moral bioenhancement program. The transparency of belief is a matter of being able to introspect one's own mental states, not of knowing where they come from. Thus, if a moral bioenhancement program manipulates participants' beliefs without also manipulating the evidence upon which they are based, it will run into the reflection problem whether or not the program is overt and compulsory like water fluoridation programs or it is entirely covert, though the problematic pattern of reasoning may be more likely if a person knows one is a participant in a program. But what if a program also changes a person's evidence for the new moral belief or changes the belief by way of changing a person's evidence?

Evidence or reasons are mental states of a person that have content; they are mental states that represent the world a certain way. Evidence only exists in the minds of believers. There is no evidence such that no one has it. Whether a belief is justified for a person depends on whether he or she has evidence that supports it. Thus, a belief can be weakly justified for a person with weak evidence, and the same belief strongly justified for a person who has lots of evidence. Justification is a matter of the relation between the content of a belief and the possessed evidence that supports the holding of that belief.¹⁸ In the

¹⁸ This is an internalist conception of evidence and justification, but such a conception has widespread appeal and is widely adopted by epistemologists.

background here is the old problem with the structure of justification, a detailed discussion of which is far afield from the present purposes. But if we think of evidence or reasons as nothing other than different beliefs, then the problem discussed above will just be pushed back a level, resulting in the famous skeptical regress of justification, because those beliefs which are supposed to justify will also need to be justified by changes in evidence, and so on. If justifiers for beliefs are beliefs themselves, then the justifiers need to be justified by something, and then those justifiers need to be justified by additional beliefs. To avoid this regress, the change in evidence that the moral bioenhancement program would need to induce is a change in non-doxastic evidence.

Appealing to non-doxastic states to provide foundational justification for beliefs is by now a common view of how we come to have justification to hold our beliefs. The types of states that are typically thought to be non-doxastic evidence for beliefs are perceptual states, such as visual representations of external world objects. Such states are not beliefs, but they are representational and, according to many, provide one justification to believe the contents of that representation. For example, if a person visually represents a red square, that representation, though not a belief, provides evidence for the belief that there is a

red square and so, in the absence of a countervailing reasons,¹⁹ one has justification to hold that belief.²⁰

Some authors extend this view to moral beliefs. According to this view, our moral intuitions, or moral “seemings,” are not themselves beliefs but are nevertheless representational.²¹ And so just like a visual representation of a red square can provide one justification for the belief that there is a red square, a moral intuition that torture is wrong can provide one justification for the moral belief that torture is wrong. So, what if a moral bioenhancement program induces both the desired moral beliefs as well as the moral intuitions that may provide participants with the justification for those beliefs?

If a moral bioenhancement program were to induce both the desired moral beliefs as well as—or even by way of—the non-doxastic evidence that supports them, then the reflection problem would be resolved, even the most introspective participant would find that his or her new moral beliefs are supported by the appropriate evidence. The sort of moral bioenhancement mentioned above that manipulates participants’ beliefs by making them more sensitive to moral reasons

¹⁹ Such as the belief that one’s vision is unreliable or reasons that the square is not red.

²⁰ Pryor, J. The Skeptic and the Dogmatist. *Noûs* 2000; 34(4): 517–549.

²¹ Huemer, M. 2008. *Ethical Intuitionism*. London. Palgrave Macmillan.

or better at moral reasoning would be this sort of program.²² The reflection problem can therefore be resolved by adopting a moral epistemology according to which non-doxastic mental states provide justification for our moral beliefs and the requirement that the program induces changes in these states. Recall, however, that the reflection problem is problem regardless of whether the program is overt and compulsory or covert. If the resolution of the problem requires adopting such a moral epistemology, then it's necessary to consider this resolution in the context of programs of both types of transparency. Doing so, however, introduces two additional epistemological problems.

Suppose again that the program is overt and compulsory but that it is manipulating both the desired moral beliefs and the evidence which supports them. Participants will be aware of the fact that their moral beliefs and moral intuitions, or whatever other non-doxastic states, are being manipulated. The first problem is that knowing that one's evidence, or one's moral intuitions, are being artificially manipulated is like knowing that the contents of one's visual perception are being artificially manipulated. Knowing that one's moral intuition that torture is wrong is a result of a manipulative moral bioenhancement program is like knowing that one's visual representation of a red square is the result of an artificial perceptual manipulation program. In the case of perception, this is a straightforward example of visual hallucination. One can have justification to

²² Manipulations that are mere cognitive enhancements would avoid the reflection problem, but not the problems that are detailed below.

believe the contents of visual hallucinations, but not if one knows one is hallucinating, even if the visual representations are accurate. Similarly, in an overt and compulsory moral bioenhancement program that manipulates someone's moral beliefs and his or her moral intuitions, participants would know that their moral intuitions and the resulting moral beliefs are the result of moral hallucination, which would defeat the justification they have to hold those beliefs. This is the moral hallucination problem.

Ignoring the moral hallucination problem, the second problem is that it is unlikely that participants will trust the source of these manipulations. Call this the trustworthiness problem. The source of the trustworthiness problem is, like the reflection problem, a matter of what participants are aware. Most moderately intelligent adults are aware that, or of have a degree of belief that, there are no moral experts and that there is widespread moral disagreement. Whether or not there are moral experts or that moral disagreement is widespread are irrelevant, as that doesn't change the fact that these attitudes are widespread in most democratic societies.²³ Both of these attitudes undermine trust in the source of the moral beliefs and the supporting evidence. That is, participants are likely to doubt that

²³ The important point here is not whether there are moral experts or whether there is moral disagreement, but that the belief that there are no moral experts and the belief that there is moral disagreement are both widespread. Probably, in fact, there are moral experts and there is widespread moral agreement, but this is irrelevant to the point.

the administrators of the program are people who are in a position to assert moral truths and induce moral beliefs in others, unless of course the participants agree with the administrators, in which case the program will do nothing to them. But for the people who disagree with the administrators—the people whose beliefs and evidence would ultimately be manipulated by the program—they will likely doubt that the administrators are in a position to override their own moral beliefs. In short, they will think that the new moral beliefs are from an unreliable source. And this mistrust defeats the justification participants have to hold their new moral beliefs.

Such is not the case with water fluoridation or vaccination programs. For each of these, there are established experts that institute or administer the program. These experts include dentists, physicians, public health officials, and scientists. Society places a great deal of trust in these people. So, when dentists and scientists and public health officials testify that water fluoridation is a significant benefit to society and this testimony is based on science, people typically trust that such is the case. But there are no similar prevailing attitudes towards moral recommendations.²⁴ This is just to say that there are no accepted moral experts. Thus, participants in the program will know that their moral

²⁴ With the exception of perhaps clergy and religious texts in a very small number of communities. Sparrow claims that it is unlikely for there to ever be agreement (2014b)

intuitions and moral beliefs are being manipulated and have reason to doubt the reliability of their source, and so their beliefs will fail to be justified for them.

To summarize, if a content-oriented moral bioenhancement program doesn't manipulate a participants' evidence along with their moral beliefs, then the program will run into the reflection problem and the participants will fail to have justification to hold the manipulated beliefs, rendering the program ineffective. But if the program does manipulate participants' evidence, and the program is overt and compulsory, the program will run into both the moral hallucination problem as well as the trustworthiness problem, and participants will fail to have justification to hold the manipulated beliefs, rendering the program ineffective. To avoid all three of these problems, a content-oriented moral bioenhancement program must therefore manipulate participants' non-doxastic evidence for their moral beliefs and do so covertly. This is the only way a content-oriented moral bioenhancement program can be effective.²⁵

CONTENT-FREE MORAL BIOENHANCEMENT

²⁵ A content-oriented moral bioenhancement program may be effective if the population is totally unreflective and introspects very little upon their own mental states. This would essentially be just like if the program were covert. But depending on people who are totally unreflective to act reliably and predictably on induced mental states is also unlikely to be effective.

A content-free moral bioenhancement program is a program that manipulates participants' states that lack content, such as a person's motivations, affective states, or emotions or dispositions.²⁶ Because they are content-free, the states themselves are not the appropriate sorts of things that can be justified for a person. But that doesn't mean that there are not important epistemological implications for a program that manipulates them.

I am skeptical of the possibility of a completely content-free moral bioenhancement program, as a person's motivation and his or her degree of belief seem to be interdependent. A change in a person's motivation may typically result in a change in the person's degree of belief that the action is what one ought do, and a change in a person's degree of belief that the action is what one ought do may typically result in a change in the person's motivation to perform the act.²⁷ If this is right, then a manipulation of a person's motivation will usually result in a change in the person's degree of belief. And a manipulation that changes a person's degree of belief is a content-oriented manipulation.

But let us suppose that it is possible to manipulate a person's moral motivations without also manipulating the degrees of his or her moral beliefs. Even when the manipulations are completely content-free, the trustworthiness problem is an issue.

²⁶ If these are in fact non-representational, though it's common to think that they are representational and so content-oriented.

²⁷ This is certainly the case if internalism about moral motivation is true.

Suppose that a participant believes that he or she ought to stop eating meat. Also suppose that one is motivated to stop eating meat. Because the participant knows that he or she is being manipulated by a moral bioenhancement program, he or she believes that the motivation to stop eating meat is the result of one's motivations being manipulated. Finally, suppose that because the participant believes that there are no moral experts, he or she mistrusts the source of the manipulations, believing instead that the source of the motivations is in no position to dictate a person's moral motivations.

My suggestion is that this mistrust is an epistemic virus that infects a person's moral psychology. The participant's mistrust of the source of the motivation will cause him or her to doubt whether his or her motivation to stop eating meat is a motivation to do what one ought to do. That is, the doubt of the source of the motivation to stop eating meat will cause one to doubt whether the cessation of meat eating is actually the morally appropriate action. It may even cause him or her to be motivated to refrain from stopping to eat meat. His or her doubt of the rightness of the prescribed action will then cause him or her to doubt the belief itself, thereby defeating the justification he or she has to hold it. The mistrust and doubt transmits through the path of least resistance, which goes all the way to the moral belief itself. And if the moral belief is unjustified for the person, then he or she is unlikely to perform it. More generally, the mistrust of the source of the motivation will make participants apprehensive of performing the action, making them less likely to perform it had they not been so mistrustful of the source.

To make clear the sort of mistrust I am suggesting, consider a person who is trying to quit smoking cigarettes. One is motivated to smoke every morning when he or she gets out of bed. But one knows that the source of the motivation is not one's reasons or evidence or beliefs about whether one ought to smoke. Rather, one knows that the source of the motivation is one's addiction. And because one knows that the source of motivation is one's addiction rather than one's judgments, beliefs, evidence, reasoning, and so on, one is likely to mistrust the motivation itself, thinking to oneself, "I'm not *really* motivated to smoke, it's just my addiction." However, if one had no idea about addiction and what it can do to a person's psychology, then one would likely have no mistrust of his or her motivation to smoke. I'm suggesting something similar would result from knowing that one's motivations are being manipulated as a result of the moral bioenhancement program. Because one has that knowledge, he or she would mistrust the moral motivations that he or she has, thinking to themselves, "I'm not really motivated to not eat meat, it's just the bioenhancement program." Couple this with the belief that the source of the motivation has no moral standing to do so, and the mistrust gets in. But if they didn't have that knowledge, they wouldn't have that mistrust.

The trustworthiness problem for content-free moral bioenhancement programs that are overt and compulsory stems from two facts: that participants believe that their motivations are being manipulated and that the people manipulating the motivations are in no position to do so. The first fact is by stipulation. The second fact is from the observation that the belief that there are

no moral experts is widespread and and that such a belief would undermine trust in those performing the manipulation. Thus, the way to avoid the trustworthiness problem is to discharge one of these two facts.

One might object by claiming that I am incorrect to state that the belief that there are no moral experts is widespread. But in a community whose members are told that they will be participants in a moral bioenhancement program, it is likely that the belief will become pervasive. It is not plausible that there would be widespread agreement about the moral authority of those administering the moral bioenhancement program, at least in diverse democratic societies. Perhaps there would be a great deal of trust in the administrators of the program in highly homogenous societies, but these are presumably not the societies that would be instituting the program in the first place. So, I don't think avoiding the trustworthiness problem by claiming that there would be trust in the source of the moral motivations will succeed.

The other strategy is to avoid the trustworthiness problem by discharging the first fact, that participants believe their moral motivations are being manipulated. The way to discharge this fact is to make it so they don't have that belief, which requires not disclosing to them that they are participants in such a program. In other words, the way to avoid the trustworthiness problem in a content-free moral bioenhancement program is to make the program covert.

THE RECOMMENDATION FROM THE EPISTEMOLOGY OF MORAL
BIOENHANCEMENT

I have argued that the epistemology of moral bioenhancement recommends that moral bioenhancement programs be covert for them to induce the desired moral behavior in enough people to meet the pressing moral demands that necessitate the institution of the program. No matter whether the moral bioenhancement program is content-free or content-oriented, there are epistemological problems the only solution to which is to make the program covert. There may be other reasons to recommend a covert program, and there may be reasons to think that no such program is ethically acceptable. I anticipate some would shudder at the thought of instituting a moral bioenhancement program covertly, as doing so may violate a person's autonomy and for this reason alone some may conclude that such a program is immoral. I don't find this line of reasoning compelling, as I think that if the expected utility of a moral bioenhancement program is great enough, the preservation of personal autonomy is irrelevant. But I have addressed none of these claims. All I have argued is that for a program to be effective, it must be covert.

Performing moral acts usually requires at least that one have justification to hold a particular moral belief (which requires evidence for the belief), the motivation to act in accordance with that belief (whatever the source), and an additional belief that performing the act will help in satisfying the prescriptions of the moral belief. There are lots of ways people can fail to act morally. I have merely highlighted one way in which participants in a moral bioenhancement program will likely fail to perform the actions the program prescribes.

There may also be other reasons that either content-oriented or content-free programs will be ineffective or, all things considered, have negative consequences. For example, it may be that the most straightforward way to institute a moral bioenhancement program is to increase participants' motivation to act on their pre-existing moral beliefs. The problem with this is that it requires that participants generally hold the right moral beliefs. If they end up holding the wrong moral beliefs, greater motivation to act upon them may engender undesirable behavior. After all, a rising tide lifts all ships, and if everyone is motivated more strongly to act on their beliefs, the result may be that participants end up acting on their more vicious moral beliefs.

ACKNOWLEDGMENTS

I wish to thank two anonymous reviewers for helpful comments on an earlier draft.