Neurofeedback-Based Moral Enhancement and the Notion of Morality

Koji Tachibana

ANNALS of the University of Bucharest
Philosophy Series

Vol. LXVI, no. 2, 2017
pp. 25–41.

ANALELE UNIVERSITATII BUCURESTI

NEUROFEEDBACK-BASED MORAL ENHANCEMENT AND THE NOTION OF MORALITY[1]

KOJI TACHIBANA[2]

Abstract

Some skeptics question the very possibility of moral bioenhancement by arguing that if we lack a widely acceptable notion of morality, we will not be able to accept the use of a biotechnological technique as a tool for moral bioenhancement. I will examine this skepticism and argue that the assessment of moral bioenhancement does not require such a notion of morality. In particular, I will demonstrate that this skepticism can be neutralized in the case of recent neurofeedback techniques. This goal will be accomplished in four steps. First, I will draw an outline of the skepticism against the possibility of moral bioenhancement and point out that a long-lasting dispute among moral philosophers nourishes this skepticism. Second, I will survey recent neurofeedback techniques and outline their three features: the variety of the target human faculties, such as emotion, cognition, and behavior; the flexibility or personalizability of the target brain state; and the nonclinical application of neurofeedback techniques. Third, I will argue that, by virtue of these three unique features, neurofeedback techniques can be a tool for moral bioenhancement without adopting any specific notion of morality. Fourth, I will examine the advantages and threats that neurofeedback-based moral enhancement may have. Finally, I will conclude that neurofeedback-based moral enhancement can become a new and promising tool for moral bioenhancement and requires further ethical investigations on its unique features.

Keywords: neurofeedback, real-time fMRI, decoded neurofeedback, moral bioenhancement, notion of morality.

---

[1] I would like to thank Roger Crisp for his many useful comments on an earlier version of the paper. I would also thank to the anonymous reviewers and the editors for their helpful comments and supports that contributed to improving the final version of the paper. This study is supported by the Japan Society for the Promotion of Science (JSPS) Overseas Research Fellowships and the JSPS KAKENHI (Grant Number 17K13318).

[2] Faculty of Letters, Kumamoto University, Japan. Email: < koji-t@kumamoto-u.ac.jp >.

1. Skepticism about Moral Bioenhancement

There is a group of criticisms regarding the very possibility of moral bioenhancement, which can be summarized and labeled as the skepticism based on the vagueness of the notion of morality. For example, John Shook (2012, 5) introduces its basic form as follows: "[t]he absence of a consensus upon the mechanisms of morality could prevent any agreement that a proposed moral enhancer could really be enhancing morality, whatever else it may be doing," although Shook does not commit himself to this criticism. A variation can be found in Dorothee Horstkötter (2016, 123), who, pointing out that advocates of moral bioenhancement do "not pay attention to questions of content," claims that "[l]eaving out moral values, however, renders one unable to distinguish between kind people made kinder and racist people rendered more racist."

These criticisms indicate that we will not be able to admit the use of a biotechnological technique as a tool for moral bioenhancement if we lack a widely acceptable notion of morality, namely, what it is like to be/become moral. Furthermore, this criticism can also be stated in a more powerful way by combining it with another argument that we can hardly provide a widely acceptable notion of morality in the first place. Committing himself to this argument, Robert Sparrow (2014, 22) claims that "[i]f we are going to start giving people drugs to make them more moral, we had better know what it is for someone to be more moral. But then who are advocates of moral enhancement to say that they know the answer to this question?" This strengthened criticism is a sort of skepticism about the very possibility of moral bioenhancement: this skepticism indicates that it will be nonsense to claim that a proposed biotechnological intervention is a tool for moral bioenhancement if no one can give a widely acceptable notion of morality in the first place.

Those who are skeptical about moral bioenhancement have a reason for presuming that it is practically difficult to frame such a notion of morality. As the history of moral philosophy testifies, the problem of what it is like to be moral has been one of the most long-lasting and unresolved disputes among philosophers. For example, it varies among the three major moral theories: Kantian deontology characterizes it as

the good will to obey duties and excludes the value of the consequences of actions (Kant 1788); utilitarianism describes it from the viewpoint of the consequences of actions and neglects the goodness or badness of the will that produces them (Bentham 1789); and Aristotelian virtue ethics highlights human ethical character that is prior to both the will and the consequences (Aristotle 1998). Although these incommensurable characterizations are merely one example, history demonstrates that no one has ever established a comprehensive moral theory that integrates these and other moral theories for hundreds – or thousands – of years by providing a widely acceptable notion of morality. This fact is certainly evidence that we should take seriously the practical (although not conceptual) difficulty of framing such a notion of morality. This kind of skepticism about moral bioenhancement is, therefore, not merely the claim that the absence of a shared understanding of such a notion *in praesenti* makes it difficult to advocate for the existence of moral bioenhancement based on a certain type of biomedical technique. Rather, being fueled by a historical fact that no one has ever succeeded in framing it, it claims that it is the absence *ad posterum* that makes it difficult to persuasively argue the possibility of moral bioenhancement.

A noble path to examine the possibility of moral bioenhancement without appealing directly to the notion of morality can be seen in a recent continuous debate between Ingmar Persson and Julian Savulescu (2016), and John Harris (2016) (see also Harris and Savulescu [2015]). In response to Harris's question, "What exactly makes something a moral enhancement?" (Harris 2016, 269), Persson and Savulescu take up various topics concerning human morality, such as reasoning, freedom, compulsoriness, motivation, emotion, and the weakness of the will, and examine whether a biotechnological intervention manages these moral topics. This path would be a right and noble one to follow when considering whether the intervention can be entitled moral bioenhancement because these multi-faceted considerations will gradually reveal what sort of biotechnology can be a tool for moral bioenhancement.

In contrast to that approach, this study aims at considering the possibility of moral bioenhancement through a different path. It will demonstrate that, despite any skepticism, the possibility of moral

enhancement can be assessed without appealing to any widely acceptable notion of morality.

2. Features of Recent Neurofeedback Techniques

A neurofeedback technique is a kind of biofeedback technique that enables a subject to regulate his/her emotion, cognition, and/or behavior by adjusting his/her own brain states with the use of a real-time representation of his/her brain activities. This technique has been studied since the development of the electroencephalograph (EEG) device in the 1950s. Although EEG-based neurofeedback training is conducted for the therapy of mental disorders, such as epilepsy and autism spectrum disorder, this technique hardly became mainstream in the disciplines of medicine and neuroscience because doctors and neuroscientists evaded the EEG-based neurofeedback research, partly because its low spatial resolution left it unable to fully comprehend the neural mechanism of such disorders (Robbins 2008)[3]. In the 1990s, the development of functional magnetic resonance imaging (fMRI) advanced the research on neural mechanisms with its high spatial resolution, but was not useful for neurofeedback research because of its weak temporal resolution (ranging from a 3–16 second delay). Neurofeedback-based clinical research thus has to wait for new devices that meet both conditions, namely, high temporal resolution and high spatial resolution.

Recently, technological and engineering developments have resulted in two kinds of new tools that satisfy these two conditions at once. The first tool is real-time fMRI (rtfMRI). As for the examples of neurofeedback research using rtfMRI, deCharms and colleagues (2005) report that patients who had chronic pain could reduce their pain perception with rtfMRI-based neurofeedback training. On the basis of earlier literature that the activity of the rostral anterior cingulate cortex (rACC) plays a crucial role in chronic pain, they set rACC as the target

[3] Marzbani and colleagues (2016) give a review on recent EEG-based neurofeedback research.

region of interest (ROI) for their experiment. Patients were asked to watch a monitor on which they could see a flame, the size of which corresponded with rtfMRI BOLD signals of their rACC. Through controlling the size of the flame on the monitor, the patients unknowingly manipulated the activity of the rACC. After the course of the training, the participants' pain perceptions were observed to be reduced. RtfMRI-based neurofeedback training research has been rapidly increasing (deCharms 2008; Weiskopf 2012). Researchers have reported its efficacy against other disorders, such as Parkinson's disease (Subramanian *et al.* 2011), (major) depression (Linden *et al.* 2012; Young *et al.* 2017), schizophrenia (Ruiz *et al.* 2013), contamination anxiety (Scheinost *et al.* 2013), alcoholism (Cox *et al.* 2016), eating disorders (Bartholdy *et al.* 2013), and pain modulation (deCharms *et al.* 2005; Chapin *et al.* 2012).

The other tool is decoded neurofeedback (DecNef) and functional connectivity-based neurofeedback (FCNef), which are different from the usual rtfMRI in that they are based on a statistical analysis of the multi-volumetric pixel (voxel) pattern of all of the brain activity, whereas the rtfMRI approach uses ROI-based analysis (Norman *et al.* 2006; LaConte 2011). The voxel-based analysis makes it possible to investigate microscopic spatial patterns of brain activity in a more specific way: DecNef and FCNef can be region-free neurofeedback training methods (Watanabe, Sasaki, Shibata *et al.* 2017)[4]. Shibata and colleagues (2011) demonstrate that subjects can acquire the perceptual ability to distinguish the orientation in a Gabor patch with DecNef training in controlling the size of a green circle and without showing any orientation figures. Currently, this technique also enables participants to learn associating orientation with color in early visual areas (Amano *et al.* 2016) and to change their facial preferences (Shibata *et al.* 2016), only through training in controlling the size of the metaphorical circle. As Shibata (2012, 1187) puts it, DecNef is a new technique that "enables us to test cause-and-effect relationships between neural activation in a

[4] As for FCNef, a further cautious assessment will be required since FCNef is characterized as being based on resting-state functional connectivity (rs-fc), but the rs-fc is under dispute about its reliability (Noble *et al.* 2017; Winder *et al.* 2017).

target brain area and changes in perception, cognition, and behavior." Recent FCNef research includes a study on reducing the severity of depression (Yamada *et al.* 2017) and a study on strengthening the cognitive performance of normal participants (Yamashita *et al.* 2017). Along with rtfMRI-based neurofeedback research, DecNef and FCNef research target their therapeutic application, especially in therapy for mental disorders (Watanabe, Sasaki, Shibata *et al.* 2017).

Unlike other possible tools for bioenhancement, such as pharmaceutical, surgical, and genetic tools, the neurofeedback techniques have various unique features including safeness, noninvasiveness, minimum side-effects, and compatibility with other approaches (Scharnowski and Weiskopf 2015; Nakazawa, Yamamoto, Tachibana *et al.* 2016; Sitaram *et al.* 2017). Among them, the following three features are important for our purpose here. The first is the variety of the target human faculties. As the aforementioned studies show, neurofeedback training can alter various human faculties, such as emotional, cognitive, and behavioral (abbreviated as ECB) states. The second is the flexibility or personalizability of the target brain state. In the case of training a certain type of emotion, for example, the target brain state can be calibrated and adjusted per participant to fit with individual specificity. In other words, the neurofeedback technique can provide personalized or tailor-made training; as Scharnowski and Weiskopf (2015, 125) state, neurofeedback can "be personalized to the individual needs of each participant in terms of the training objectives, and the target brain areas/networks." The third feature is the nonclinical application of neurofeedback techniques. Neurofeedback studies have already been conducted on healthy individuals as well as those with mental disorders.

3. Neurofeedback Training and Moral Enhancement

As Gordijn and Buyx (2010) foresee, the focus of neural engineering in general will shift slowly but certainly from therapy to enhancement. As for neurofeedback techniques, some neuroethicists suggest that the advancement of neurofeedback research will make neurofeedback-based moral enhancement possible (Jotterand and

Giordano 2015; Nakazawa, Yamamoto, Tachibana *et al.* 2016). This possibility, however, similar to the other possible types of moral bioenhancement, is a matter of controversy among its advocates (Kabasenche 2016; Maslen and Savulescu 2016) and skeptics (Crutchfield 2016; Horstkötter 2016). The skepticism against the possibility of moral bioenhancement based on the lack of a shared understanding of morality plays an important role in this controversy. As outlined in Section 1, a longstanding dispute among moral philosophers on this issue encourages this skepticism.

However, it is worth noting that the dispute among moral philosophers sheds another light on human morality that the skeptics might overlook. That is, those different moral theories highlight and represent different constituents of human morality, each of which we take seriously in our daily lives: we appreciate, not exclusive-disjunctively but conjunctively, the will, behavior, and character (emotion and cognition), which correspond approximately to ECB[5]. In short, the goodness and badness – or in general the state – of ECB, or the performances of ECB, determine the human moral states. It would be nonsense to assume that those major moral theories do not reflect any aspect of our common-sense understanding of morality since, as Aristotle rightly puts it, well-known and long-lasting opinions represent aspects of truth (Aristotle 1998, VII1). Therefore, if we have no way to solve the dispute nor to reach a widely acceptable notion of morality, it might mean that our notion of morality is not so sophisticated or coherent. Nevertheless, it remains valid to argue that ECB performance plays a decisive role for our moral state or performance. Thus, the dispute among moral philosophers does not merely have the negative claim that the skeptics discern; it also permits us to draw a positive lesson that the apparently mutually-exclusive characterizations of morality between such moral theories, approximately their different emphases on ECB, highlight different aspects of our common-sense understanding of morality. Since, as mentioned in section 2, a

[5] The will can be understood as a sort of rational decision that is based on emotion and cognition and occurs prior to behavior.

neurofeedback technique can change ECB performance, such a technique can be a tool for changing moral state.

One might still argue that, even if a change in the ECB performance through neurofeedback training can bring about a change in moral performance, the change in moral performance does not necessarily imply that the (result of) change is a moral *enhancement* because it leaves open whether the change improves or worsens one's moral performance. Here we see the problem of the notion of morality. This problem would not be solved by appealing to the definition of enhancement. For example, enhancement is defined as "any change in the biology or psychology of a person which increases the chances of leading to a good life in the relevant set of circumstances" (Savulescu *et al.* 2011, 7). Therefore, following this persuasive definition of human enhancement, the change in moral performance through neurofeedback training will rightly be moral bioenhancement if it leads to a good life in moral contexts, whereas it will be moral *deterioration* if it disturbs a good life in such contexts. However, this would not be helpful for our purpose because it remains uncertain what a *good* life is. Rather, this just sends the problem back to another and probably more complex problem of the vagueness of the notion of a good life. Thus, even if enhancement is properly defined and neurofeedback training is rightly said as a tool for changing human morality, it cannot be regarded as a tool for *enhancing* morality without embodying the notion of morality (and/or a good life). And what is worse, whatever notion of morality is adopted, it would hardly be accepted widely because its rival theories would not voice their approvals with the notion. This nourishes the skepticism against the very possibility of moral bioenhancement.

Neurofeedback techniques can neutralize this skepticism on their use for moral bioenhancement. To elaborate upon this point, let us consider the two difficulties in the efficacy assessment of neurofeedback training on morality. The first concerns the measurement of personalized interventions that is also well known in psychotherapeutics. The psychotherapeutic procedure for mental disorders is usually so personalized that randomized controlled trials (RCTs) do not evaluate its efficacy well (Cooper 2007, chap. 9; cf. Young *et al.* 2017). Given that neurofeedback-based moral bioenhancement will be performed in a

personalized way, the results may not be assessed well with RCTs or other standard statistical procedures that are used to validate efficacy in typical clinical trials (Nakazawa, Yamamoto, Tachibana *et al.* 2016). The second concerns the vagueness of the notion of morality or the problem of making morality explicit. Even if we have a good gauge for measuring the efficacy of training, we may not be able to identify the result as moral improvement until we articulate a widely acceptable notion of morality.

Each of these does show the downside of the possibility of neurofeedback-based moral bioenhancement. However, as a minus multiplied by a minus makes a plus, these two negative factors, being combined with each other, bring a fruitful result, namely, cancelling out the difficulties in the efficacy assessment of neurofeedback training on morality. First, an RCT is a statistical procedure to avoid subjective biases and provide an objective assessment of the efficacy of a scientific (usually medical) intervention with setting objective endpoints, blinding, and so on. During the course of an RCT, the moral aspects of medical intervention yield a subjective bias because people have different thoughts on what is "moral" (Jadad 1998). Therefore, in the case of assessing the efficacy of moral – not medical – interventions, such as moral bioenhancement, vagueness is unavoidably observed regarding what the outcome and the endpoint objective are. This is an inevitable situation in the case of RCT-based efficacy assessment on moral intervention since there has never been a consensus or a widely-acceptable notion of morality. To begin with, accordingly, as for a moral intervention, an RCT is not an appropriate way of evaluating the efficacy of moral interventions. This is neither the fault of RCTs nor of moral evaluation. Moral evaluation just does not lend itself well to RCTs.[6]

Second, against the skeptics, the lack of a shared understanding of morality does not necessarily imply that neurofeedback training cannot be a tool for moral bioenhancement or, in other words, a tool for providing morally good results. Many activities can be identified as good providers without having any widely acceptable notion of the goodness of products. For example, presumably, people have never

[6] Here, the second feature neutralizes the first.

reached any widely acceptable notion of what is considered the most stylish hairstyle or clothing. Nevertheless, each of us can give (confidently or not) our own opinion on it. Although our opinions can be and are divided, this does not stop us from agreeing on what constitutes a good hairdresser or tailor: we can agree that a hairdresser who can cut a customer's hair as he/she likes is a good hairdresser and a tailor who can provide what his/her customer wants is a good tailor. The goodness of the hairdresser or tailor depends on whether he/she can provide what his/her customer wants, and not on whether he/she has a widely acceptable notion of stylish hair or clothing. A customer's friends may deny (or even laugh at) the customer's aesthetic sense because they have different aesthetic senses. Nevertheless, the customer and his/her friends will be able to reach an agreement that the hairdresser or tailor is good if the hairdresser or tailor provides exactly what the customer wants. Generally speaking, in the case of evaluating the goodness of a personalized or tailor-made activity, the goodness involves whether the activity provides the targeted result, and not whether the provider prefers the result. The same applies to the efficacy assessment of neurofeedback training on morality insofar as, being different from other techniques, such as pharmaceuticals, this technique is highly personalized or tailor-made. Schematically speaking, the more accurately a neurofeedback training provides what each party wants, the more easily Kantians and utilitarians will reach an agreement that neurofeedback training is a good tool for enhancing morality, without agreeing on a universal notion of morality.[7]

Flexibility or personalizability is an important factor in assessing the social feasibility of moral enhancement (Harris 2016; Shook 2012; Sparrow 2014). All the proposed tools for moral enhancement thus far fail to be moral enhancements because they do not have sufficient flexibility to realize a variety of moral states. For example, a pharmaceutical that may enhance one's deontic volitional power will be criticized from a utilitarian viewpoint because it cannot enhance one's utilitarian disposition. Any single pharmaceutical can be criticized by one or another moral party insofar as it adopts different notions of

[7] Here, the first feature neutralizes the second.

morality. In contrast, in the case of neurofeedback training on morality, by virtue of its variety of target moral faculties and its highly personalized or tailor-made procedure, a single device can be used for both deontic and utilitarian enhancements or any morality that each party wants. It can even enhance human vulnerability, in which some critics of bioenhancement discern an ethical value (Parents 1995; McKenny 1998). This flexibility makes neurofeedback-based moral enhancement a unique exception of the skepticism based on the notion of morality. Thus, this very feature makes it possible to use neurofeedback training as a tool for moral bioenhancement without being committed to any specific notion of morality. Accordingly, the skepticism can be neutralized in the case of the moral bioenhancement-purposive use of a neurofeedback technique.[8]

4. Advantages of and Threats in Neurofeedback-based Moral Enhancement

Neurofeedback-based moral enhancement has several advantages in the moral bioenhancement debate. First, as discussed thus far, it avoids the skepticism based on the vagueness of the notion of morality. Recent neurofeedback techniques, such as rtfMRI and DecNef, do not require a widely acceptable notion of morality to claim their eligibility to be used for moral bioenhancement. This means that advocates of moral bioenhancement can send the cumbersome problem of the notion of morality back to moral philosophers.

Second, it saves moral diversity. Moral bioenhancement in general is sometimes blamed for its inflexibility that may lead to moral perfectionism and moral uniformity, which is not compatible with moral diversity, the value of which is socially appreciated (Sparrow 2014; Wiseman 2014). Neurofeedback-based moral bioenhancement would

[8] Although this argument highlights the possibility of moral neuroenhancement, the same argument will be applicable for that of cognitive enhancement because the former includes the latter, but not *vice versa* (Tachibana 2009; see also Carter and Gordon 2015).

avoid this criticism because its variety of target human faculties and the flexibility or personalizability of the target moral state neither support moral perfectionism, nor bring moral uniformity, but instead save (or even advance) the diversity.

Third, since the emergence of this technique, moral bioenhancement has become a real possibility for two reasons. The first is that neurofeedback research has already been conducted on healthy people as well as people with mental disorders. On one hand, some neuroscientists have shown that rtfMRI neurofeedback is effective for healthy subjects to self-regulate their positive and negative emotions, such as happiness, disgust, sadness, tenderness, and affection (Moll *et al.* 2014; Sitaram *et al.* 2011; Zotev *et al.* 2011). These studies can be instances of neurofeedback-based enhancement on moral emotions. On the other hand, DecNef studies on the association of orientation with color (Amano *et al.* 2016) and the regulation of facial preferences (Shibata *et al.* 2016) may realize neurofeedback-based moral enhancement to eliminate discrimination based on skin-color or appearance. Furthermore, as for other possible aspects of morality, such as empathy and altruistic behavior, once the neural network is statistically identified through leveling the brain states of moral exemplars who have such an aspect of morality, then neurofeedback-based moral enhancement will be performed by setting the statistically standardized brain state or network as the target figure. The second reason is the legal and economic features of neurofeedback-based moral enhancement. In the case of pharmaceutical moral bioenhancement, drug patents will have a disincentive for pharmaceutical moral bioenhancement for legal and economic reasons, such as intellectual property (Socaciu and Uszkai 2015). In contrast, neurofeedback-based moral bioenhancement will not concern any intellectual-property-related issues because it only uses fMRI and brain data as the target figure. This system will have long-term economic efficiency. Such scientific progress and legal or economic features, as well as the features mentioned above, will promote the realization of neurofeedback-based moral bioenhancement. By virtue of these advantages, this technique will certainly advance the moral bioenhancement debate.

However, we should not also fail to recognize the threats that such advantages may bring. Related to the first and second advantages, we need to examine whether this technique might intensify the differences among people and finally lead to moral relativism. A man of deontic temptation will not want to become utilitarian but more deontic, whereas a man of utilitarian temptation will not want to become deontic but more utilitarian. Generally speaking, people are inclined to enhance what they appreciate, regardless of whether it is physical, cognitive, or moral enhancement (cf. Gyngell and Easteal 2015). As this technique can help people realize anything they want, it may influence them to become increasingly different from one another. If neurofeedback-based moral enhancement enhances parochial morality and incites moral relativism, then it may neither be socially feasible nor endorsed by the family of nations. The problem of the lack of a shared understanding of morality reappears here, but in a different way. As mentioned previously, neurofeedback-based moral enhancement can send the cumbersome problem back to moral philosophers. However, this does not mean disregarding the problem. Rather, it returns the problem to the proper place where it should be investigated – the field of traditional moral philosophy. If neurofeedback-based moral enhancement incites moral relativism as an actual threat to global society, then the problem of the notion of morality becomes a pressing issue not only for us but also for moral philosophers.

Related to the third advantage, the threat of moralization and medicalization needs to be considered. For decades, conceptual and political efforts have been devoted, on one hand, to remove from the psychological and psychiatric approaches and treatments of mental disorders the tendency to moralize mental disorders and its undesired consequences, and, on the other hand, to remove from our society the tendency to medicalize morally different people and their cultures. However, these two different issues – the treatment of mental disorders and the moral correction – may overlap with each other by the emergence of neurofeedback-based moral enhancement. From antiquity to the present, the relationship between them has conceptually been considered both positively (Jaeger 1957; Martin 2006) and negatively (Foucault 1961; Szasz 2010; Wiseman 2014). However, the threat that the emergence of

neurofeedback-based moral enhancement may bring is not merely conceptual, but also practical. The fact that the same device with the same procedure can be successfully used for controlling both mental and moral states may lead people to identify mental disorders with moral deficits and assume that moral differences are the subject of treatment.

5. Conclusion

Recent neurofeedback techniques will enable a new type of moral bioenhancement that does not require any specific notion of morality. On one hand, this will advance the moral bioenhancement debate. However, on the other hand, it will introduce new neuroethical issues, some of which may become real threats to our society. The discussion that this study provides is not decisive for demonstrating neurofeedback-based moral bioenhancement to be socially feasible. Further neuroethical examination is required.

REFERENCES

Amano, K., K. Shibata, M. Kawato, Y. Sasaki, and T. Watanabe (2016). "Learning to associate orientation with color in early visual areas by associative decoded fMRI neurofeedback." *Current Biology*, 26(14): 1861-6.

Aristotle (1998). *Nicomachean Ethics*. Translated by W. D. Ross, revised by J. L. Ackrill, and J. O. Urmson. Oxford, UK: Oxford University Press.

Bartholdy, S., P. Musiat, I. C. Campbell, and U. Schmidt (2013). "The potential of neurofeedback in the treatment of eating disorders: A review of the literature." *European Eating Disorders Review*, 21: 456-63.

Bentham, J. (1789). *An Introduction to the Principles of Morals and Legislation*. London: T. Payne.

Carter, J., and E. Gordon (2015). "On cognitive and moral enhancement: A reply to Savulescu and Persson." *Bioethics*, 29(3): 153-61.

Chapin, H., E. Bagarinao, and S. Mackey (2012). "Real-time fMRI applied to pain management." *Neuroscience Letter*, 520(2): 174-81.

deCharms, R. C. (2008). "Applications of real-time fMRI." *Nature Review of Neuroscience*, 9(9): 720-9.

deCharms, R. C., F. Maeda, G. H. Glover, D. Ludlow, J. M. Pauly, D. Soneji, J. D. E. Gabrieli, and S. C. Mackey (2005). "Control over brain activation and pain learned by using real-time functional MRI." *Proceedings of the National Academy of Sciences*, 102(51): 18626-31.

Cooper, R. (2007). *Psychiatry and Philosophy of Science*. Stocksfield, UK: Acumen.

Cox, R. W., L. Subramanian, D. Linden, M. Lührs, R. McNamara, R. Playle, K. Hood, G. Watson, J. Whittaker, R. Sakhuja, N. Ihssen (2016). "Neurofeedback training for alcohol dependence versus treatment as usual: study protocol for a randomized controlled trial." *Trials*, 17: 480.

Crutchfield, P. (2016). "Decoded neurofeedback is unlikely to enhance moral capacities." *American Journal of Bioethics: Neuroscience*, 7(2): 125-6.

Foucault, M. (1961). *Histoire de la folie à l'âge classique*. Paris: Plon.

Gordijn, B. and A. M. Buyx (2010). "Neural engineering." In *Scientific and Philosophical Perspectives in Neuroethics*, J. Giordano, and B. Gordijn (eds.), 283–301. New York, NY: Cambridge University Press.

Gyngell, C. and S. Easteal (2015). "Cognitive diversity and moral enhancement." *The Cambridge Quarterly of Healthcare Ethics*, 24(1): 66-74.

Harris, J. and J. Savulescu (2015). "A debate about moral enhancement." *Cambridge Quarterly of Healthcare Ethics*, 24(1): 8-22.

Harris, J. (2016). "Moral blindness – the gift of the god machine." *Neuroethics*, 9(3): 269-73.

Horstkötter, D. (2016). "The ethics of novel neurotechnologies: Focus on research ethics and on moral values." *American Journal of Bioethics: Neuroscience*, 7(2): 123-4.

Jadad, A. R. (1998). *Randomized Controlled Trials: A User's Guide*. London, UK: BMJ Books.

Jaeger, W. (1957). "Aristotle's use of medicine as model of method in his Ethics." *The Journal of Hellenistic Studies*, 77(1): 54-61.

Jotterand, F. and J. Giordano (2015). "Real-time functional magnetic resonance Imaging-Brain-Computer interfacing in the assessment and treatment of psychopathy: Potential and challenges." In *Handbook of Neuroethics*, J. Clausen and N. Levy (eds.), 763-81. New York, NY: Springer.

Kabasenche, W. P. (2016). "Moral formation and moral enhancement." *American Journal of Bioethics: Neuroscience*, 7(2): 130-1.

Kant, I. (1788). *Kritik der praktischen Vernunft*. Riga: J. F. Hartknoch.

LaConte, S. M. (2011). "Decoding fMRI brain states in real-time." *Neuroimage*, 56(2): 440-54.

Linden, D. E., I. Habes, S. J. Johnston, S. Linden, R. Tatineni, L. Subramanian, B. Sorger, D. Healy, and R. Goebel (2012). "Real-time self-regulation of emotion networks in patients with depression." *PLoS One*, 7(6): e38115.

Martin, M. (2006). *From Morality to Mental Health: Virtue and Vice in a Therapeutic Culture*. Oxford, UK: Oxford University Press.

Marzbani, H., H. R. Marateb, and M. Mansourian (2016). "Neurofeedback: A comprehensive review on system design, methodology and clinical applications." *Basic and Clinical Neuroscience*, 7(2): 143-58.

Maslen, H. and J. Savulescu (2016). "Neurofeedback for moral enhancement: Irreversibility, freedom, and advantages over drugs." *American Journal of Bioethics: Neuroscience*, 7(2): 120-2.

McKenny, G. P. (1998). "Enhancements and the ethical significance of vulnerability." In: *Enhancing Human Traits: Ethical and social implications*, E. Parens (ed.), 222-37. Washington D.C.: Georgetown University Press.

Moll, J., J. H. Weingartner, P. Bado, R. Basilio, J. R. Sato, B. R. Melo, I. E. Bramati, R. de Oliveira-Souza, and R. Zahn (2014). "Voluntary enhancement of neural signatures of affiliative emotion using fMRI neurofeedback." *PLoS One*, 9(5): e97343.

Nakazawa, E., K. Yamamoto, K. Tachibana, T. Soichiro, Y. Takimoto, and A. Akabayashi (2016). "Ethics of decoded neurofeedback in clinical research, treatment, and moral enhancement." *American Journal of Bioethics: Neuroscience*, 7(2): 110-7.

Noble, S., M. N. Spann, F. Tokoglu, X. Shen, R. T. Constable, and D. Scheinost (2017). "Influences on the Test–retest reliability of functional connectivity MRI and its relationship with behavioral utility." *Cerebral Cortex*, 27(11): 5415-29.

Norman, K. A., S. M. Polyn, G. J. Detre, and J. V. Haxby (2006). "Beyond mind-reading: Multi-voxel pattern analysis of fMRI data." *Trends in Cognitive Sciences*, 10(9): 424-30.

Parens, E. (1995). "The goodness of fragility: On the prospect of genetic technologies aimed at the enhancement of human capacities." *Kennedy Institute of Ethics Journal*, 5(2): 141-53.

Persson, I. and J. Savulescu. (2016). "Moral bioenhancement, freedom and reason." *Neuroethics*, 9(3): 263-8.

Robbins, J. (2008). *A Symphony in the Brain: A Newly Revised and Expanded Edition*. New York, NY: Grover Press.

Ruiz S., S. Lee, S. R. Soekadar, A. Caria, R. Veit, T. Kircher, N. Birbaumer, and R. Sitaram (2013). "Acquired self-control of insula cortex modulates emotion recognition and brain network connectivity in schizophrenia." *Human Brain Mapping*, 34(1): 200-12.

Savulescu, J., A. Sandberg, and G. Kahane (2011). "Well-Being and Enhancement." In *Enhancing Human Capacities*, J. Savulescu, R. Meulen and G. Kahane (eds.), 1-18. Oxford, UK: Wiley-Blackwell.

Scheinost, D., T. Stoica, J. Saksa, X. Papademetris, R. T. Constable, C. Pittenger, and M. Hampson (2013). "Orbitofrontal cortex neurofeedback produces lasting changes in contamination anxiety and resting-state connectivity." *Translational Psychiatry*, 30(3): e250.

Scharnowski, F. and N. Weiskopf (2015). "Cognitive enhancement through real-time fMRI neurofeedback." *Current Opinion in Behavioral Sciences*, 4: 122-7.

Shibata, K. (2012). "A new neuroscientific approach using decoded neurofeedback (DecNef)." *Clinical Neurology*, 52: 1185-7 (in Japanese.)

Shibata K., T. Watanabe, M. Kawato, and Y. Sasaki (2016). "Differential activation patterns in the same brain region led to opposite emotional states." *PLoS Biology*, 14(9): e1002546. doi:10.1371/journal. pbio.1002546.

Shibata, K., T. Watanabe, Y. Sasaki, and M. Kawato (2011). "Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation." *Science*, 334(6061): 1413-5.

Shook, J. R. (2012). "Neuroethics and the possible types of moral enhancement." *American Journal of Bioethics: Neuroscience*, 3(4): 3-14.

Sitaram, R., S. Lee, S. Ruiz, M. Rana, R. Veit, and N. Birbaumer (2011). "Real-time support vector classification and feedback of multiple emotional brain states." *Neuroimage*, 56(2): 753-65.

Sitaram, R., T. Ros, L. Stoeckel, S. Haller, F. Scharnowski, J. Lewis-Peacock, N. Weiskopf, M. L. Blefari, M. Rana, E. Oblak, N. Birbaumer, and J. Sulzer (2017). "Closed-loop brain training: The science of neurofeedback." *Nature Review Neuroscience*, 18: 86-100.

Socaciu, E. M. and R. Uszkai (2015). "How drug patents might lead to disincentives for moral bioenhancement." *Ethics in Biology, Engineering and Medicine: An International Journal*, 6(1-2): 125-37.

Sparrow, R. 2014. "Egalitarianism and moral bioenhancement." *American Journal of Bioethics*, 14(4): 20-8.

Subramanian, L., J. V. Hindle, S. Johnston, M. V. Roberts, M. Husain, R. Goebel, and D. Linden (2011). "Real-time functional magnetic resonance imaging neurofeedback for treatment of Parkinson's disease." *Journal of Neuroscience*, 31(45): 16309-17.

Szasz, T. (2010). *The Myth of Mental Illness: Foundations of a Theory of Personal Conduct*. New York, NY: Harper Perennial.

Tachibana, K. (2009). "Can moral enhancement be a subject of neuroethics? Neuroethics and the governance of brain science." *Departmental Bulletin of Department of History and Philosophy of Science, University of Tokyo*, 11: 1-35 (in Japanese.)

Watanabe, T., Y. Sasaki, K. Shibata, and M. Kawato (2017). "Advances in fMRI real-time neurofeedback." *Trends in Cognitive Neurosciences*, < http://dx.doi.org/10.1016/j.tics. > 2017.09.010.

Weiskopf, N. (2012). "Real-time fMRI and its application to neurofeedback." *Neuroimage*, 62(2): 682-92.

Winder, A. T., C. Echagarruga, Q. Zhang, and P. J. Drew (2017). "Weak correlations between hemodynamic signals and ongoing neural activity during the resting state." *Nature Neuroscience*, doi:10.1038/s41593-017-0007-y

Wiseman, H. (2014). "Moral enhancement – 'hard' and 'soft' forms." *The American Journal of Bioethics*, 14(4): 48-9.

Yamada, T., R. Hashimoto, N. Yahata, N. Ichikawa, Y. Yoshihara, Y. Okamoto, N. Kato, H. Takahashi, and M. Kawato (2017). "Resting-state functional connectivity-based biomarkers and functional MRI-based neurofeedback for psychiatric disorders: a challenge for developing theranostic biomarkers." *International Journal of Neuropsychopharmacology*, 20(10): 769-81.

Yamashita, A., S. Hayasaka, M. Kawato, and H. Imamizu (2017). "Connectivity neurofeedback training can differentially change functional connectivity and cognitive performance." *Cerebral Cortex*, 27(10): 4960-70.

Young, K. D., G. J. Siegle, V. Zotev, R. Phillips, M. Misaki, H. Yuan, W. C. Drevets, and J. Bodurka (2017). "Randomized clinical trial of real-time fMRI amygdala neurofeedback for major depressive disorder: Effects on symptoms and autobiographical memory recall." *The American Journal of Psychiatry*, doi:10.1176/appi.ajp.2017.16060637

Zotev, V., F. Krueger, R. Phillips, R. P. Alvarez, W. Kyle Simmons, P. Bellgowan, W. C. Drevets, and J. Bodurka (2011). "Self-regulation of amygdala activation using real-time fMRI neurofeedback." *PLos One*, 6(9): e24522.