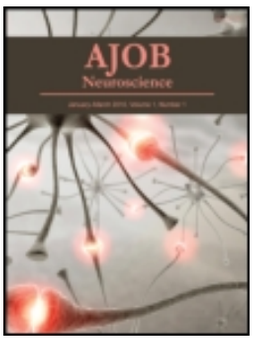

AJOB Neuroscience

ISSN: 2150-7740 (Print) 2150-7759 (Online) Journal homepage: http://www.tandfonline.com/loi/uabn20

# Moral Enhancement? Acknowledging Limitations of Neurotechnology and Morality

John R. Shook & James Giordano

**To cite this article:** John R. Shook & James Giordano (2016) Moral Enhancement? Acknowledging Limitations of Neurotechnology and Morality, AJOB Neuroscience, 7:2, 118-120, DOI: 10.1080/21507740.2016.1188178

**To link to this article:** http://dx.doi.org/10.1080/21507740.2016.1188178

Published online: 18 Jul 2016.

*AJOB Neuroscience*, 7(2): 118–133, 2016

ISSN: 2150-7740 print / 2150-7759 online
DOI: 10.1080/21507740.2016.1188178



**Open Peer Commentaries**

# Moral Enhancement? Acknowledging Limitations of Neurotechnology and Morality

**John R. Shook,** University at Buffalo
**James Giordano,** Georgetown University Medical Center

Nakazawa, Yamamoto, Tachibana, and colleagues (2016) offer their hopes and concerns about employing real-time functional magnetic resonance imaging (rt-fMRI)-based neurofeedback to treat mental disorders and to enhance moral cognition. Their recognition of limitations to these techniques is welcome, given exaggerated claims that often characterize discussions of neurotechnologically derived enhancements. Discerning the extent to which any type of enhancement is achievable depends on a number of factors, ranging from the neurological to the sociological (Shook, Galvagni, and Giordano 2014; Shook and Giordano 2016). The meaning of "moral enhancement" at minimum depends on the approach(es) used to alter brain function, the experimental protocols for attaining envisioned goals, and the shared understandings of experimenters, subjects, and society about the significance of those goals for morality. Before attempting to judge the moral worth or ethical status of any alteration in brain function, we must first be prepared to explain and justify how some neurological modification could even be classified as an improvement upon a person's morality. How such a verifiable classification is accomplished will then provide information required for evaluating whether some putative improvement to morality is both authentic and ethical.

As the authors describe them, protocols utilizing decoded neurofeedback rt-fMRI appear to elicit validly effective results. Such protocols (i.e., to alter subjects' ability to more "accurately" make moral judgments, which have been preset as targets) do seem feasible. The authors are optimistic that their technique can simultaneously affect multiple brain regions, even though they admit that brain networks involved in moral cognition are as yet only tentatively identified and poorly understood. This is not a huge obstacle, since recent studies are discerning details about brain regions networked in various types of moral judgment (Cushman 2013; Avram et al. 2014). In fact, rt-fMRI enables subject-by-subject inquiry into the efficacy and durability of adjustments to moral cognition, with all the attendant risks openly declared in advance, which may contribute to further advances in the field. Nevertheless, the neurological processes underlying social cognition in general and moral cognition in particular won't be adequately understood anytime soon. We still must address key questions about the capabilities, limits, and value of the neurotechnology—and method(s)—used (Giordano 2015). Indeed, as the authors note, there are several limitations, mainly arising from diffuse effects upon complex networks that happen to be involved with moral cognition and other modes of cognition as well.

We contend that there are (at least) four additional limitations—not coincidentally involving social factors—that are significant when considering adjustments to a person's morality.

First, a subject can produce different moral judgments without anyone, including experimenters, understanding which components of moral cognition have been adjusted and why those adjustments caused differing moral judgments. Subjects would be unable to say why they think differently about moral matters, even in the ordinary terms of folk moral psychology; this might be disorienting and disconcerting. Confusion could be reduced if protocols included identifying alterations to affective, motivational, valuational, or reasoning processes during procedures (see, e.g., the work of Moll et al. [2014] and Sherwood et al. [2016]).

Second, a subject undergoing this technique would be encouraged by experimenters to adjust moral judgment away from what initially seems intuitive and "right enough," toward judgments that can't, by definition, seem quite right to the subject. After all, if the subject took some alternative judgment to be "just as morally good" then the

Address correspondence to John R. Shook, Department of Philosophy and Graduate School of Education, University at Buffalo, 135 Park Hall, Buffalo NY, USA. E-mail: jrshook@buffalo.edu

goal or outcome wouldn't be "moral adjustment." How would sincere subject compliance be guaranteed if and when subjects feel like they must make morally variant judgments during the procedure, and anticipate becoming morally "different" or even immoral if the procedure works for them? Compliance could be enhanced by assuring subjects that the goal of adjustment is only to make moral judgment more consistently and strongly like judgments they already subjectively regard as moral.

Third, a subject undergoing this technique might successfully transition to a new manner of moralizing and take this new condition to be morally right. As the authors note, new moral habits may drift away from expected standards, or even fade away, over time. Subjects might demand continued treatments in order to "stay moral." What happens if those treatments are not available? This is a real problem because there is little (or no) guarantee that such drift would revert to a subject's original morality. This technique cannot be labeled as "reversible" quite yet. Sustaining and/or reversing such interventions could pose a problematic issue, unless reliably long-term techniques were developed.

Fourth, a subject undergoing this technique might be informed that the procedure is needed to improve one's moral judgment in order to correct a mental disease or disorder. Does this offer a way to treat "moral pathologies"? More importantly, do such clinical-sounding classifications (which we would hesitate to make) appear to establish "moral enhancement" as primarily about therapeutic rehabilitations or reformations? Setting up external standards of morality rather than subjective standards places experimental protocols in very different territory, where sincere and voluntary compliance surely cannot be presumed. Furthermore, pathologizing many modes of deficient or defective morality can look like the start of a slippery slope, leading toward social condemnation of "undesirable" moralistic stances.

Most fundamentally, we urge closer scrutiny into treating these experimental protocols as genuinely moral enhancements. Neurological improvements are not automatically ethical enhancements (Shook and Giordano 2016). In this light, we raise four concerns that arise in parallel with the four key questions about experimental protocols from the first half of this commentary.

Our first concern attends to the authors' suggestion that reducing differences in the default mode network between "healthy" people and patients with mental disorders offers a route toward moral enhancement. This may be so, but current justifications of moral improvement of individuals focus upon their deficiencies in folk moral psychology: A person may not care enough about others, or not be nice enough, and so on. Suppose that this experimental technique enables a person to behave "better." There is no promise that the subject will introspectively grasp why, and subjects may even report little to no sensed change in their moral psychology: They may not feel like they care more, want to share more, and so forth. After all, terms of folk moral psychology merely represent, and not necessarily in any closely corresponding way, hypothesized tokens of neurological function and structure(s). Would we trust this procedure to enhance morality if we can't explain why a subject has become morally better? Improved behavior is not automatically moral behavior, and it may be far from ethically commendable conduct (Shook 2012). Something else may be causing observed behavioral changes and evoking other unanticipated and/or undesirable side effects.

Our second concern pursues the idea that compliance could be strengthened if subjects believe that their preexisting moral habits were (only) to be strengthened and intensified. How would society regard that mode of moral enhancement? Moral pluralism is both a psychological fact and a tolerated reality in society, to certain limits (Graham et al. 2011). Would society be comfortable with a large portion becoming more "conservative" and another large portion becoming more "liberal" (to use Haidt's terms)? Enhancement does not display an obvious unitary and unified directionality. Additionally, what if some people want to calibrate their morality toward some chosen ethical exemplar? Why be just a little more conservative, if you could acquire the moral judgment of a right-wing media celebrity? Let's not be naive—there will be those who would seek these "boutique" sorts of enhancement.

Our third concern proceeds from our observation that moral remediation and rehabilitation may become technologically feasible at the expense of becoming socially dangerous. Those who call for moral enhancement for many people must worry about who will classify—and be classified as—the morally "healthy" or "unhealthy." Who among us is really so morally healthy? On the other hand, if we permit people to enhance in their preferred directions, does that only encourage the moral tribalization of humanity? The world needs less tribalization, not more, but homogeneity doesn't seem right, either. Perhaps this neurotechnology, along with others, could be used to discern a statistically average "morally normal" brain by averaging together $n$ number of individuals' default mode networks. Yet which individuals, from which cultures, would be selected for that sample? That "normalized" moral brain wouldn't even seem quite moral, or appear to be morally vacillating or inconsistent, from the stance of those who expect deontic, utilitarian, or virtue ethics to dominate a person's moral psychology.

Our fourth and related concern centers upon how much ethical responsibility and control must be exercised when this technique is applied on human subjects, no matter how far advanced it may become. To alter a person's internal moral sense and judgment is to manipulate something at the core of who we are as responsible agents and personal selves. That we already do this for children (fairly well) and moral deviants (not so well) does not in any way diminish the significance of, and responsibility for, this ambitious endeavor. To experiment with long-term moral judgment and moral conduct is to undertake nothing less than social reengineering on a grand scale. The

dystopian literature about political impositions of a "normalized" morality on populations prompts ethical unease if not outrage, and rightly so. May we never surrender that ethical wisdom to neurotechnical prowess or moral expediency. ■

## REFERENCES


Avram, M., K. Hennig-Fast, Y. Bao, et al. 2014. Neural correlates of moral judgments in first- and third-person perspectives: Implications for neuroethics and beyond. *BMC Neuroscience* 15:39. doi:10.1186/1471-2202-15-39.

Cushman, F. 2013. Action, outcome, and value: A dual-system framework for morality. *Personality and Social Psychology Review* 17 (3): 273–92. doi:10.1177/1088868313495594.

Giordano, J. 2015. A preparatory neuroethical approach to assessing developments in neurotechnology. *AMA Journal of Ethics* 17(1): 56–61.

Graham, J., B. Nosek, J. Haidt, R. Iyer, S. Koleva, and P. Ditto. 2011. Mapping the moral domain. *Journal of Personality and Social Psychology* 101(2): 366–85.

Moll, J., J. Weingartner, P. Bado, et al. 2014. Voluntary enhancement of neural signatures of affiliative emotion using fMRI neurofeedback. *PLoS ONE* 9(5): e97343. doi:10.1371/journal.pone.0097343.

Nakazawa, E., K. Yamamoto, K. Tachibana, S. Toda, Y. Takimoto, and A. Akabayashi. 2016. Ethics of decoded neurofeedback in clinical research, treatment, and moral enhancement. *AJOB Neuroscience* 7(2): 110–117.

Sherwood, M., J. Kane, M. Weisend, and J. Parker. 2016. Enhanced control of dorsolateral prefrontal cortex neurophysiology with real-time functional magnetic resonance imaging (rt-fMRI) neurofeedback training and working memory practice. *NeuroImage* 124(Pt A): 214–23. doi:10.1016/j.neuroimage.2015.08.074

Shook, J. R. 2012. Neuroethics and the possible types of moral enhancement. *AJOB-Neuroscience* 3(4): 3–14.

Shook, J. R., L. Galvagni, and J. Giordano. 2014. Cognitive enhancement kept within contexts: Neuroethics and informed public policy. *Frontiers of Systems Neuroscience* 8:228.

Shook, J. R., and J. Giordano. 2016. Neuroethics beyond normal: Performance enablement and self-transformative technologies. *Cambridge Quarterly of Health Care Ethics* 25(1): 121–40.




# Neurofeedback for Moral Enhancement: Irreversibility, Freedom, and Advantages Over Drugs

**Hannah Maslen,** University of Oxford
**Julian Savulescu,** University of Oxford

Nakazawa and colleagues (2016) examine potential therapeutic applications of decoded neurofeedback for the treatment of psychiatric conditions such as depression, and developmental disorders. Decoded neurofeedback, they argue, is particularly promising in this regard, since it can enable individuals to observe a representation of their brain activity in real time. Consequently, individuals can train themselves to intentionally adjust their brain activity, ultimately in the absence of the visual representation. Nakazawa and colleagues further hypothesize that decoded neurofeedback techniques could be used for moral enhancement, if individuals were able to train themselves to adjust their brain states to those conducive to moral behavior, such as the brain state correlated with compassion. This, they argue, would be a particularly appealing form of moral enhancement, since modulation of brain states could be "personalized, or tailor made" to the individual's beliefs about how to live morally.

The article makes an important contribution in drawing the attention of neuroethicists to the prospect that neurofeedback could enable a greater degree of control over mental phenomena such as emotions or strong desires, which sometimes frustrate our ability to act in line with our moral commitments, or prudentially. Although Nakazawa and colleagues are optimistic about the way in which the personalization of neurofeedback preserves moral pluralism, they suggest that there are significant ethical concerns relating to irreversibility, safety, and efficacy. Irreversibility in particular, they argue, is potentially problematic in its implications for freedom, since agents are rendered unable to alter themselves if their moral beliefs change. This, they argue, is in contrast to pharmaceuticals,

Address correspondence to Hannah Maslen, The Oxford Uehiro Centre for Practical Ethics, University of Oxford, Suite 8, Littlegate House, 16/17 St Ebbe's Street, Oxford, OX1 1PT, United Kingdom. E-mail: Hannah.maslen@philosophy.ox.ac.uk