

OPC on Zawadzki & Adamczyk, “Personality and Authenticity in Light of the Memory-Modifying Potential of Optogenetics”

Alexandre Erler

alexandre.erler@philosophy.oxon.org

This is the final, pre-publication draft of an Open Peer Commentary published in AJOB Neuroscience, 12:1 (2021), 40-42, DOI: [10.1080/21507740.2020.1866108](https://doi.org/10.1080/21507740.2020.1866108). Please cite the published version. (Note: This research was supported by a GRF grant (14607918) from the Research Grants Council (RGC) of Hong Kong.)

Optogenetic memory modification and the many facets of authenticity

Zawadzki and Adamczyk (forthcoming) offer a rich and enlightening overview of the unique potential of optogenetics as a memory modification technology (MMT), particularly with regards to the unsettling prospect of (reversibly) erasing unpleasant autobiographical memories – including pathological, but perhaps also “just painful” ones. In this commentary, I would like to raise two issues regarding their discussion of the implications of optogenetic memory modification for personality and authenticity. The first one has to do with the scope of that discussion, and the nature of the concept of authenticity. The second one concerns the exact reason why, in the authors’ view, potential changes in a subject’s values resulting from such a procedure should be viewed as ethically problematic; as well as what the implications are at the policy level.

Zawadzki and Adamczyk use the coherentist model advocated by Pugh, Maslen and Savulescu (2017) as the basis of their assessment of the potential impact on authenticity from optogenetic memory erasure. As they indicate, reasoning from that model, concerns about authenticity mostly arise in cases where the procedure affects the subject’s *values*, or alters some of her traits in ways that conflict with these values (Zawadzki and Adamczyk, forthcoming, p. 12). Such an approach treats the notion of authenticity as very much akin to that of *autonomy*: its guiding idea is that living authentically means living in accordance with our “true” values and desires, however these are to be spelt out exactly.

It is hard to dispute that autonomy is a very important ethical consideration, and one that is often associated, if not equated with the concept of authenticity. Nevertheless, I believe we should recognize that considerations of autonomy do not exhaust legitimate concerns about authenticity, including in the present context. The key reason for this, I contend, is that authenticity is an ambiguous, or *multifaceted* concept: there is more than one legitimate, plausible conception of it. As an analogy, consider the concept of consciousness, described by Ned Block as a “mongrel concept” (Block, 1995). Arguably, “consciousness” can refer to different, though related, entities (phenomenal consciousness, access consciousness, etc.), so that it would be misguided to ask which of these reflects the “correct” understanding of consciousness.

I maintain that something similar is likely true of authenticity. While the concept does seem to have some kind of essence or “core”, namely the idea of being in some sense “true”, “truthful” or “real”, there are multiple ways, plausible and well-established, to develop that core, yielding a plurality of legitimate conceptions of authenticity. Authenticity as autonomy is one such conception, yet there are others. Consider authenticity as *truthful living*: on this conception, living authentically means living a life based on a mostly accurate apprehension of one’s own personal circumstances, including the significant aspects of one’s past. The

value of authenticity thus understood has been emphasized by authors like Carl Elliott (Elliott, 1998), the President's Council on Bioethics (2003), and myself (Erlor, 2011). To my knowledge, no decisive argument has so far been provided for favouring one particular conception of authenticity over all plausible alternatives (simply appealing to our own intuitions about conceptual fit may be a common practice, but is unlikely to convince those who do not already share them). In light of this, I argue that the best approach is to adopt a form of *pluralism* about authenticity.¹

Several plausible conceptions of authenticity seem relevant to an ethical analysis of optogenetic memory erasure. Besides authenticity as autonomy, it may be a necessary consequence of such a procedure that it will negatively impact authenticity as truthfulness, insofar as it involves depriving the person of significant information about her past. (How problematic that might be remains of course a matter of debate.) Furthermore, it seems that it might sometimes promote (or at least not affect) authenticity in one sense, while impeding it in another. Someone could conceivably decide, in a manner fully consonant with her values, to erase an important autobiographical memory and go on to live a less truthful, yet happier life. Preserving authenticity as understood by Pugh and colleagues would then not yet entail that the person's authenticity had not been impacted in any other relevant sense. That is not to say that any discussion of such a case should necessarily aim to cover all the relevant senses of authenticity. Rather, it just means that we should not use an inappropriately monistic approach to the concept as a reason to pre-emptively dismiss other conceptions of it that could have normative significance.

Secondly, even if we assume Pugh and colleagues' coherentist and autonomy-focused view of authenticity as a basis for discussion, I believe the need remains to further explain why exactly, in Elizabeth's case as discussed by Zawadzki and Adamczyk, a significant change in her values (and possibly other aspects of her psychology) should elicit ethical concern.² In particular, it seems important to distinguish between different considerations that might be relevant in this context. The authors' mention of the possibility that Elizabeth might eventually "resign from her anti-bullying organization in order to...find a more profitable career" after erasing her memories of victimization (Zawadzki and Adamczyk, forthcoming, p. 14), suggests that the concern here is about *self-corruption*: Elizabeth's values might change for the worse, leading her to become less concerned about helping others, and more focused on the pursuit of material gain for herself. However, while such a concern may well apply in certain cases (though presumably not all), it is arguably distinct from the authors' claim that the changes in Elizabeth's psychology "must be traceable over the diachronic process of intelligible rational change" (*ibid.*).

¹ A pluralistic approach to the concept has previously been defended by Erik Parens (e.g. Parens, 2014). That said, the pluralism I advocate goes further than that of Parens, and does not fundamentally rely on a dichotomy between two approaches or frameworks ("gratitude" vs. "creativity" in Parens's words, or self-discovery vs. self-creation, as mentioned by Zawadzki and Adamczyk). Authenticity as truthful living, for instance, does not seem to fall neatly on either side of such dichotomies.

² Assuming such a change would occur, which, as the authors outline towards the end of their article, may not be so likely after all, for instance given the possible persistence of relevant semantic self-knowledge that could help sustain Elizabeth's original values. Given the important role other people would play in this regard, this in turn raises interesting questions that I do not have the space to tackle here: e.g. do others have anything like a duty not to remind a person of what she had chosen to forget (and if so, how is it to be balanced against their own rights, such as freedom of speech)? Or on the contrary, as the authors briefly consider (Zawadzki and Adamczyk, forthcoming, p. 14), do they have a duty to help protect the person's authenticity by making sure she does not *completely* forget about her self-defining past experiences?

Zawadzki and Adamczyk emphasize “stability of values” as a key condition of the possibility such a process (*ibid.*). This suggests that the fundamental concern, on their view, might relate to changes in Elizabeth’s values that she may not have anticipated and consented to beforehand: such changes may not be intelligible in light of her original, highest-priority values (or perhaps “meta-values”), which might not condone the loss of her commitment to fight bullying for the sake of gaining “peace of mind”. However, on this line of thought, if her original higher-order values did condone such a tradeoff, the change in her attitude towards anti-bullying activism following the optogenetic procedure need not raise any concerns about authenticity. The issue then essentially becomes one of informed consent.

One would like to know whether this accurately reflects the authors’ main line of argument, and if so, what they take to be its practical implications. Were cases to arise in which threats to authenticity of the kind they describe seemed to present themselves, would this be sufficient grounds for restricting access to optogenetic memory modification? Or alternatively, out of respect for what has been termed people’s right to mental self-determination (Bublitz and Merkel, 2014), should any future providers of optogenetic MMTs simply work to ensure that each user had been properly informed about any such risks? And should it make a difference whether the procedure is or not therapeutic in nature?

The important ethical conversation highlighted by Zawadzki and Adamczyk should undoubtedly continue. Virtually all of the considerations I have touched on here strike me as potentially significant. One lesson to be drawn from these brief reflections might be that a switch from ambiguous talk of authenticity to the use of more precise ethical concepts (such as autonomy, truthful living, self-corruption, etc.) could benefit the debate. While I see no persuasive reason to demand that people avoid the language of authenticity altogether when expressing concerns about memory modification, making sure we disambiguate that concept when it is used can help highlight the distinct and potentially conflicting values being invoked, and minimize the risk of talking past one another.

REFERENCES:

- BLOCK, N. 1995. On a Confusion About the Function of Consciousness. *Behavioral and Brain Sciences*, 18, 227-47.
- BUBLITZ, J. C. & MERKEL, R. 2014. Crimes Against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination. *Crim Law and Philos*, 8, 51-77.
- ELLIOTT, C. 1998. The Tyranny of Happiness: Ethics and Cosmetic Psychopharmacology. In: PARENS, E. (ed.) *Enhancing Human Traits: Ethical and Social Implications*. Washington, DC: Georgetown University Press, 177-86.
- ERLER, A. 2011. Does Memory Modification Threaten Our Authenticity? *Neuroethics*, 4 (3), 235-49.
- PARENS, E. 2014. *Shaping our Selves : On Technology, Flourishing, and a Habit of Thinking*, Oxford, Oxford University Press.
- PUGH, J., MASLEN, H. & SAVULESCU, J. 2017. Deep Brain Stimulation, Authenticity and Value. *Camb Q Healthc Ethics*, 26, 640-657.
- THE PRESIDENT'S COUNCIL ON BIOETHICS (U.S.) 2003. *Beyond Therapy : Biotechnology and the Pursuit of Happiness*, New York, ReganBooks.
- ZAWADZKI, P. & ADAMCZYK, A. K. Forthcoming. Personality and Authenticity in Light of the Memory-Modifying Potential of Optogenetics. *AJOB Neuroscience*.