Optogenetic Manipulation of Maladaptive Memory–New Challenges or New Solutions for Personal Authenticity?

Elsey, J.W.B.

Link to publication

Optogenetic manipulation of maladaptive memory –
new challenges or new solutions for personal authenticity?

James W. B. Elsey, PhD.

University of Amsterdam, Department of Clinical Psychology

Our memories in many ways define us. Our understanding of who we are and what we have been through, what we believe in and what matters to us, are all tied to traces of our pasts, etched into our brains. Hence, the prospect that optogenetic manipulation of memory might enable the targeted silencing of our most defining memories rightly raises concerns that, in seeking self-enhancement, we may stumble into self-negation. Zawadzki and Adamczyk consider the many ways in which they believe optogenetics may be both uniquely powerful and appealing, but also uniquely risky, as a means of modifying human memory, with a particular focus on challenges to personal authenticity. I raise two challenges to the perspective of Zawadzki and Adamczyk that may sharpen our understanding of the implications of optogenetics. Firstly, I highlight some gaps in our knowledge of the capabilities of optogenetics that, depending on how they are resolved, may render optogenetics more or less unique among memory-modifying technologies. Secondly, I discuss some neglected capabilities of optogenetics that may pose substantially less risk to personal authenticity than memory silencing or erasure.

The desirability of optogenetic memory manipulation will depend on how targeted and specific it really can be. Much of the animal research on optogenetic manipulation of memory has involved tagging of memories during memory formation (Ramirez, Tonegawa, and Liu 2014). Less work has focused on targeting memories that have already been formed, and especially those formed long ago. Similar approaches to

tagging during formation could be used, but note that in animals we can do this in a highly controlled environment that may even replicate the conditions of memory formation. Tagging of meaningful human memories may be substantially more problematic, with all sorts of related memory traces being reactivated simultaneously, running the risk that viral targeting spreads beyond the desired target. Indeed, given the many levels of meta-representation that may be built up out of a disturbing memory, and the wide range of brain regions different forms of representation may be stored in (Roy et al. 2020), it becomes challenging to know even what *should* be targeted. It will require considerably more work before we know the translational feasibility of highly selective targeting for the kinds of memory Zawadzki and Adamczyk discuss. If these problems cannot be surmounted, then the appeal of optogenetics for silencing memories will likely be minimal.

Nevertheless, we continue to be stunned by the sophistication of tools and precision with which researchers show themselves able to manipulate memory in non-human animals. It may be inevitable that precision targeting of even complex human memory ultimately becomes possible. When we have this level of sophistication, is it likely that we use it to produce wholesale silencing or loss of entire episodes of our past, rather than a more selective approach that reduces suffering but does not pose such a threat to our sense of self?

Perhaps surprisingly, Zawadzki and Adamczyck suggest that it is an advantage of optogenetic procedures that they can prevent a bad memory's retrieval altogether, rather than merely remove its 'sting', as in other memory modifying technologies (MMTs). To the contrary, it is largely because of the idea that many current MMTs do not erase memories or remove declarative knowledge of our pasts that they are ethically palatable (Elsey and Kindt 2016). Moreover, this selective form of 'amnesia' is not just a random phenomenon. It points to a certain architecture of memory, in line with

theories of the distributed engram (Josselyn and Tonegawa 2020; Roy et al. 2020), that could presumably also be leveraged by optogenetic manipulation to produce less problematic and selective forms of amnesia than brute silencing of the entire memory. For example, it may be possible to target specific regions of the brain or components of the engram responsible for particular effects – such as insufferable emotional reliving – while leaving other elements of the engram intact. Indeed, this might already be occurring in some studies of optogenetic memory modification: We do not have access to the subjective experience of laboratory animals, and do not strictly know whether certain optogenetic manipulations are really experienced as a total loss of memory. Careful observation of animal behavior after pharmacological memory modification has suggested that behavioral outcomes taken to demonstrate memory loss may reflect diminished motivation produced by the manipulation (Cogan et al. 2019), and selective silencing of parts of the engram through optogenetics might likewise affect only certain forms of memory expression even among animals. In short, it is not clear that optogenetic approaches actually do remove all trace of a memory from consciousness, and if they do, it is likely that more selective approaches could be developed that only target specific engram components.

To take the discussion further, we can grant that optogenetic silencing of maladaptive memories in humans could be possible and might be experienced essentialy as a form of total memory loss or erasure. As Zawadzki and Adamczyk argue, if a person were suddenly given the ability not to better cope with or move on from some past trauma, but to entirely forget it had ever happened, this would risk the person losing something that – for better or for worse – makes them who they are. Although I suspect there are cases in which a person would gladly lose such a part of their self, or in which retaining a certain memory affords no prospect of personal development or redemption, people probably share the intuition that in most cases something of ethical significance

may be lost along with a silenced memory. But perhaps optogenetics provides ways to achieve the same goal as people who resort to memory silencing might seek, without posing a clear threat to personal authenticity.

Such a possibility is suggested by contemporary optogenetic research, and aligns with theories of the therapeutic process in humans (Brewin 2006). Many psychological disturbances can be seen as disorders of maladaptive memory (Elsey and Kindt 2017), in which the effects of past experience linger, being expressed in excessive and unpleasant emotional reactions or constant rumination/fixation. Directly diminishing the power of such maladaptive memory traces, as might be pursued through optogenetic silencing, is only one way in which their power may be challenged. An alternative method, and that which is believed to underpin extinction learning in animals (Bouton 2002) and the effects of dominant psychological therapies in humans (Brewin 2006), is the production of alternative adaptive memory traces that compete with maladaptive ones. Rather than silencing maladaptive memories, it may instead be possible to bolster adaptive representations. Several studies have demonstrated that specific neuronal ensembles can be identified that are responsible for learned reductions in maladaptive memory expression. An important role has been given to neurons in the infralimbic cortex (part of the prefrontal cortex) for reducing both threat-related (Kim et al. 2016) and reward-related responding (Villaruel et al. 2019). It has also been suggested that distinct engrams even within a single brain region – the hippocampus – may be responsible for memory expression and suppression (Lacagnina et al. 2019). These studies found that such populations of neurons could be identified and controlled optogenetically, with concomitant changes in memory expression.

If optogenetic manipulation of maladaptive memory traces in humans is possible, then it should possible to optogenetically target adaptive memory traces as well. This may be a particularly appealing approach with respect to personal authenticity. Rather

than simply preventing a memory from being recalled, bolstering a person's adaptive capacities could be experienced as more consistent with a 'natural' developmental process, in which they gain greater control over their troubling thoughts and feelings. In some senses, the person may even feel more authentic, having achieved the mastery they aspired to, but were unable to realise. While some may see this as ultimately an inauthentic achievement, one wonders if it is reasonable to restrict a person's sense of ownership to their possibly deeply troubling and in some cases even futile 'natural' struggle, the outcome of which may equally be determined by factors over which they cannot strictly claim responsibility – chance environmental occurrences and congenital factors. At the very least, this approach would seem to offer more authenticity than simply silencing a person's past.

In summary, the idea that optogenetic approaches pose unique challenges to personal authenticity is not well-founded. If optogenetic approaches develop to the level of sophistication for use in humans, I think there will also be ways that it could be used without necessarily just silencing memories. More authenticity-preserving approaches will likely be favored by users. Despite my challenges, Zawadzki and Adamczyk are right to consider whether there may be novel or unique ethical concerns with optogenetics. Thinking about routine use of optogenetic technology in humans seems almost absurd at present, yet the enthusiasm with which similarly invasive technologies such as Neuralink have been received suggests a market for such devices. Ethical implications of these technologies will not be restricted to authenticity or even memory, but will extend to all aspects of our mental lives.

Bouton, ME. 2002. "Context, Ambiguity, and Unlearning: Sources of Relapse after Behavioral Extinction." *Biological Psychiatry* 52 (10): 976–86. https://www.sciencedirect.com/science/article/pii/S0006322302015469.

Brewin, Chris R. 2006. "Understanding Cognitive Behaviour Therapy: A Retrieval

Competition Account." *Behaviour Research and Therapy* 44 (6): 765–84.

https://doi.org/10.1016/j.brat.2006.02.005.

Cogan, ES, MA Shapses, TE Robinson, and N Tronson. 2019. "Disrupting Reconsolidation:

Memory Erasure or Blunting of Emotional/Motivational Value?"

Neuropsychopharmacology 44 (2): 399–407. https://www.nature.com/articles/s41386-

018-0082-0.

Elsey, J. W. B., and M Kindt. 2017. "Tackling Maladaptive Memories through

Reconsolidation: From Neural to Clinical Science." *Neurobiology of Learning and*

Memory 142: 108–17.

https://www.sciencedirect.com/science/article/pii/S1074742717300382.

Elsey, J.W.B., and M. Kindt. 2016. "Manipulating Human Memory Through Reconsolidation:

Ethical Implications of a New Therapeutic Approach." *AJOB Neuroscience* 7 (4).

https://doi.org/10.1080/21507740.2016.1218377.

Josselyn, Sheena A, and Susumu Tonegawa. 2020. "Memory Engrams: Recalling the Past and

Imagining the Future." *Science* 367 (eaaw4325).

https://doi.org/10.1126/science.aaw4325.

Kim, HS, HY Cho, GJ Augustine, and JH Han. 2016. "Selective Control of Fear Expression

by Optogenetic Manipulation of Infralimbic Cortex after Extinction."

Neuropsychopharmacology 41 (5): 1261–73.

https://www.nature.com/articles/npp2015276.

Lacagnina, A. F, E. T Brockway, C. R Crovetti, F. Shue, M. J McCarthy, K. P Sattler, and M.

R Drew. 2019. "Distinct Hippocampal Engrams Control Extinction and Relapse of Fear

Memory." *Nature Neuroscience* 22 (5): 753–61.

https://www.nature.com/articles/s41593-019-0361-z.

Ramirez, Steve, Susumu Tonegawa, and Xu Liu. 2014. "Identification and Optogenetic

Manipulation of Memory Engrams in the Hippocampus." *Frontiers in Behavioral Neuroscience*. https://doi.org/10.3389/fnbeh.2013.00226.

Roy, Dheeraj S, Young-Gyun Park, Sachie K Ogawa, Jae H Cho, Heejin Choi, Lee Kamensky, Jared Martin, Kwanghun Chung, and Susumu Tonegawa. 2020. "Brain-Wide Mapping of Contextual Fear Memory Engram Ensembles Supports the Dispersed Engram Complex Hypothesis." *Biorxiv*. https://doi.org/10.1101/668483.

Villaruel, FR, F Lacroix, C Sanio, DW Sparks, CA Chapman, and N Chaudhri. 2019. "Optogenetic Activation of the Infralimbic Cortex Suppresses the Return of Appetitive Pavlovian-Conditioned Responding Following Extinction." *Cerebral Cortex* 28 (12): 4210–21. https://academic.oup.com/cercor/article-abstract/28/12/4210/4555262.