ORIGINAL PAPER

# The Ethics of Memory Modification: Personal Narratives, Relational Selves and Autonomy

**Przemysław Zawadzki**

**Abstract** For nearly two decades, ethicists have expressed concerns that the further development and use of memory modification technologies (MMTs)—techniques allowing to intentionally and selectively alter memories—may threaten the very foundations of who we are, our personal identity, and thus pose a threat to our well-being, or even undermine our "humaneness." This paper examines the potential ramifications of memory-modifying interventions such as changing the valence of targeted memories and selective deactivation of a particular memory as these interventions appear to be at the same time potentially both most promising clinically as well as menacing to identity. However, unlike previous works discussing the potential consequences of MMTs, this article analyzes them in the context of the narrative relational approach to personal identity and potential issues related to autonomy. I argue that such a perspective brings to light the ethical aspects and moral issues arising from the use of MMTs that have been hidden from previously adopted approaches. In particular, this perspective demonstrates how important the social context in which an individual lives is for the ethical evaluation of a given memory-modifying intervention. I conclude by suggesting that undertaking memory modifications without taking into account the social dimension of a person's life creates the risk that she will not be able to meet one of the basic human needs—the autonomous construction and maintenance of personal identity. Based on this conclusion, I offer some reflections on the permissibility and advisability of MMTs and what these considerations suggest for the future.

## Introduction

For nearly two decades, ethicists have expressed concerns that the further development and use of memory modification technologies (MMTs)—techniques allowing to intentionally and selectively alter memories—may threaten the very foundations of who we are, our personal identity, and thus pose a threat to our well-being, or even undermine our "humaneness" [1–24]. The most auspicious currently studied MMTs such as propranolol administration [25], extinction training combined with the memory update [26] or with administration of glucocorticoids [27, 28] have the potential to alter memory in a relatively safe, selective and

P. Zawadzki (✉)
Faculty of Philosophy, Institute of Philosophy, Jagiellonian University, Krakow, Poland
e-mail: przemyslaw.zawadzki@uj.edu.pl

clinically relevant manner. These memory-modifying interventions offer to alter emotional components or certain expressions of excessively emotional (fear, phobic, or traumatic) memory. In such interventions declarative memory (e.g., belief phobic system) and episodic memory (e.g., conscious awareness of the experienced trauma) remain intact [29].

However, since the persistence of episodic memories is highly relevant to posttraumatic stress disorder (PTSD)[1] and other anxiety disorders [33], it may sometimes be necessary to render amnesia of traumatic or phobic memory as this can be the only way to overcome PTSD and ultimately build healthy identity. Deactivation of memories has been realized in model organisms for decades with the use of protein synthesis inhibitors. However, while protein synthesis inhibitors constitute proof of concept, their utilization in humans is precluded due to their toxicity. Moreover, they are not as selective as one could expect of memory modification intervention that is to be used in humans—protein synthesis inhibitors impair all consolidating memories [34]. In contrast, findings of recent years have shown that the prospect of providing measures to selectively deactivate the targeted memory can be actually realized with the use of an emerging MMT—optogenetics [35–37]. Although making optogenetic system operable to control neural activity underlying the memory still involves invasive manipulations, i.e., inserting opsin genes via viral infection and implanting optical fibers, less invasive optogenetic–systems are currently under development (see [38–42] for recent advancements in making optogenetics less or even non-invasive). While it may take years to make optogenetics or optogenetic-like technology a therapeutic memory–modifying

measure,[2] if ever to be realized,[3] the fact that MMT has been shown to be effective in deactivating specific memories makes it timely and important to consider potential philosophical[4] and ethical ramifications of such an intervention among other memory-modifying interventions.

I analyze the potential ramifications of memory modifications such as changing the valence of targeted memories and selective deactivation of a particular memory as these interventions seem to be at the same time potentially both most promising clinically as well as menacing to personal identity. In contrast to previous works discussing the potential consequences of MMTs, this paper examines them in the context of the narrative relational approach to personal identity and potential issues related to autonomy. I argue that such a perspective brings to light ethical aspects and moral issues arising from the use of MMTs that have been hidden from previously adopted approaches. In particular, this perspective shows how important is the social context in which the individual is embedded for the ethical evaluation of a given memory modification intervention—as opposed to the more individualistic accounts framed around the concept of authenticity that have previously been adopted in the neuroethical literature [3, 5, 12, 13, 17, 23].

Moreover, in contrast to recent works attempting to analyze the normative implications of MMTs in the light of the concept of authenticity that would "take relational nature of personal identity seriously" [8, 14, 53], I argue, following Mackenzie and Walker

---

[1] PTSD is a psychological disorder which has dramatic consequences for a daily life of affected person. It is characterized by heightened noradrenergic signaling [30], highly disruptive traumatic memories, avoidance, and increased arousal and stress response [31]. In PTSD traumatic memories are constantly retrieved even in non-threatening (neutral) situations, triggering re-experiencing of the trauma. Re-experiencing can involve intrusive and vivid episodic memory of the traumatic experience and evoke adverse emotions associated with it [32]. Finally, PTSD can lead to the development of associated pathologies such as depression, aggression, substance abuse and high risk of suicidal behaviors [31].

[2] Although it is worth noting that the first clinical trial for neurodegenerative disease in humans has already been successfully carried out—optogenetics allowed to partially restore visual function in a patient with retinitis pigmentosa [43].

[3] There is an ongoing discussion on whether using optogenetics as a therapeutic MMT will be possible in the future [1, 44–50]. There are several threads in this debate: safety concerns [44, 47], technological obstacles to translating optogenetic techniques from animal models to humans [44, 47], and issues associated with the nature of memory, e.g., whether it will be possible to: control specific memory contents in human brains [44, 47], track complex autobiographical memories [44, 49], or selectively erase particular autobiographical memories due to their 'non-local' nature [44, 48, 50].

[4] Optogenetics has been recently discussed as a tool that may help to solve conceptual puzzles regarding the nature of memory [51, 52].

[54], that such interpretation of authenticity "blurs distinction between identity and autonomy and so obscures the central ethical issue it seeks to address." The central ethical issue underlying narratives of self-alienation concerns threats to autonomy since "a relational and narrative account of identity and autonomy can incorporate the most plausible aspects of the ethics of authenticity and can explain what is at stake in the first-person phenomenological descriptions of self-alienation" [54]. Consider the following excerpt from Leuenberger (2022) discussing her processual relational narrative account of authenticity in the context of the potential ethical implications brought about by MMTs: "By changing the constraints of who you can be, neural interventions open up the possibility for life paths that are easier and maybe happier while remaining genuine. The shy person would not have to force himself to appear outgoing when applying for a job favoring an outgoing attitude. He would just genuinely feel like being outgoing. However, this raises serious issues of conformity, which would contradict a further pillar of authenticity: that we should not bend to the will of other people to fit in but be ourselves" [14]. Similarly, Iftode, Zorila, Vica, Mihailov (2022) claim in concluding remarks of their recent discussion on the ethics of MMTs in the light of their experimental relational account of authenticity that: "However, the inescapable experimental side of authenticity makes us avoid confusing one's personal identity with a social and narrative construct that only bears the mark of social pressure and conformism. In other words, sticking to the value of authenticity we are making sure that a richness of life alternatives remains available within our society, and that our 'right to difference, variation and metamorphosis' (Deleuze, 1988, p. 106) is preserved" [8].

The above passages demonstrate that what is really at stake in the discussion concerning the ethics of relational authenticity is the issue of being autonomous: "that we should not bend to the will of other people to fit in but be ourselves" and "that a richness of life alternatives remains available within our society, and that our 'right to difference, variation and metamorphosis' is preserved" are calls for (the prospect of) autonomy. Relational accounts of authenticity share with each other that they attempt to accommodate the ideal of autonomy as one of the pillars of authenticity. However, it may not be a good idea because autonomy is a better defined and more value-laden concept—from which the concept of relational authenticity arguably derived its normativity in the first place.[5] Therefore, I agree with diagnosis of Mackenzie and Walker [54] that we should focus in the neuroethical analyzes on the concepts of identity and autonomy rather than authenticity—if we are to enrich our analyzes with a relational dimension. This perspective allows for an explicit and thorough analysis of the most salient ethical issues arising from the use of MMT.

## Narrative Relational Accounts of Personal Identity

Narrative accounts of personal identity can be classified as normative or descriptive [56]. One of the most frequently discussed narrative accounts of personal identity in neuroethical considerations is Schechtman's Narrative Self-Constitution View [12, 16, 57–62]. Schechtman's conception is normative (but see [61, pp. 101–102]) as she states that narrative is something that one must cultivate to be a person and imposes constraints on the narrative that is to be identified as the constitutive of personal identity [62, pp. 113–114]. The first of them is the articulation constraint, i.e., the thesis that "an identity-constituting narrative must be capable of local articulation" [62, pp. 114]. An identity-constituting narrative need not cover the entire life of the person, but she should be able to narrate parts that matter for explaining why she acts, believes, and feels as she does [62, pp. 114]. The articulation constraint is thought to ensure the fulfillment of common sense expectations of personal identity—that the individual's sense of diachronicity and the prospect of self-identification is secured. The second requirement, the reality constraint, indicates that the narrative must be consistent with basic observational and interpretative facts about the individual, her life and the world. One cannot simply claim that she is Vladimir Lenin as defending such an identity claim would require to contradict obvious, almost

---

[5] This view seems to be gaining ground in recent debate [49, 55].

indisputable facts about this person or even about how the world works. Thus, although in Schechtman's view personal identity is ultimately the product of a person's inner life, there are limits to the stories that a person can tell about herself to make an authoritative act of the identity self-constitution. According to Schechtman's model, these limits are dictated by others' stories regarding the person in question. In this approach, other persons' narrative perspectives are testimonies serving to verify if the reality constraint is fulfilled, thereby whether the person's autobiographical narrative has authoritativeness in constituting her own identity.

In most feminists' approaches, other persons' narratives play even more important role, that is, they co-constitute one's identity along with the first-person narrative. According to Lindemann [56, pp. 81], personal identity depends upon social recognition as "many practical identities require more that one person for their construction and maintenance." And even more radically, when expressing her relational account of identity-constitution, Lindemann [63, pp. 210] claims that humans "without other people could not be persons at all." At the same time, however, she emphasizes that persons are not only what society makes of them; persons can also defy society's expectations and decide for themselves who they ultimately will be:

Indeed, we can think of human selves as comprising two intertwined strands that are often in tension and even, in certain cultures and at certain times in history, become unbalanced because one strand takes ascendancy over the other. Call one strand "the given" and the other "the chosen." "The given" consists of our first and much of our second natures, the age and society into which we were born, the relationships with which we were encumbered at birth, the identities others impose on us, our first and maybe second language, and morality itself. "The chosen" embraces our status as agents who choose freely and act on the basis of those choices and includes our ability to reason, our free will, our autonomy, and our capacity to reflectively endorse or repudiate the considerations that bear on what we do and what we think [63, pp. 209].

Thus, according to Lindemann "the given" is not everything there is that counts when it comes to personal identity. Obviously, a person must "settle oneself" in the given in order to be able to construct narrative; however, a person is also characterized by "the chosen"—she can deliberate on what stories emphasize values, social roles, relationships, actions she cares about and choose to adopt in her self-narrative the ones that do. First-person perspective of "the chosen" and third person perspective of "the given" often align with each other. However, the crucial question to understand Lindemann's account is to ask what if these perspectives are in conflict with each other. The answer to this question determines which of them prevail—that is, which one of the two is constitutive for one's identity—and on what terms. Lindemann [56, pp. 101] proposes to resolve this question through introducing the normative requirement for the narrative that is to be identity constituting. This requirement has a form of the credibility constraint. Lindemann postulates three credibility criteria for evaluating which of the perspectives should be privileged and authoritative for identity constitution.

In contrast to Schechtman's account—for which the third-person perspective is epistemically privileged as it functions as a point of reference for the first-person story that is often characterized by self-serving biases, self-deceptiveness, blissful ignorance, and other factors distorting narrative truth-preservation—Lindemann claims that the third-personal perspective may be just as skewed as the first-personal one. Thus, the epistemic privilege of the first-person perspective cannot be justified a priori; it is not, therefore, sufficient to follow the reality constraint imposed by Schechtman's Narrative Self-Constitution View and to only ensure that our identity aligns with the view that others hold regarding ourselves. Lindemann claims that in order to determine which of these perspectives provides the better story, one should follow credibility criteria. These are: explanatory force, correlation to action, and heft. The narrative has stronger explanatory force if it fits the evidence for explaining the person's characteristics better; it conforms with the second criterion more completely if its correlation with person's actions is more fitting—not necessarily with one's intentions, conscious deliberations, goals, or beliefs; last but not least, a perspective is higher on the criterion of heft if it has a greater weight and importance regarding the pertinent characteristics of the person's autobiographical narrative. For Lindemann, therefore, an identity-constituting narrative can be identified with the most credible story told by (or about) a person.

The narrative account of Baylis [64] coincides with Lindemann's perspective in the most central theme of relational accounts of personal identity as she claims that "persons are (and can only be) dynamic complex co-creations constituted and maintained through iterative and cyclical as well as public and private performances" [64, pp. 119]; thus, in Baylis view, we are simultaneously constantly projecting ourselves through our stories and we—our selves—are the projects of identity-creations that are imposed on us by others. In contrast to Lindemann, however, for Baylis an identity-constitutive story cannot be identified with the content of the narrative that turned out to be the winner from the two (or more) competing perspectives that better fulfill credibility criteria, but rather with the content of the story that emerges after finding the "equilibrium." "Equilibrium" is "the balance between how a person sees and understands herself and how others see and understand her" [64, pp. 119]. Moreover, in contrast to Schechtman's account in which the crucial role of the fulfillment of the reality constraint by the narrative that is to be authoritative for identity self-constitution is emphasized, according to Baylis, "there is neither care nor concern with what others 'objectively' believe to be true, except insofar as others' perceptions are invitations to practise one's narrative and performance skills" [64, pp. 121].

For Baylis, a prerequisite for identity-constitution is others reception of the projected self-narrative—preferred (the story of who the person wants to be), performed (the story of who the person can be, given all constraints that are imposed on her by the self and others), or combination thereof. There are several ways in which the projected narrative can be uptaken by others. It may be: dismissed, resisted, tolerated, actively endorsed, or actively contested in whole or in part. In the case of negative reception of the self-narrative, one's response can take at least five different forms in order to satisfy the equilibrium constraint: the person may try to project her preferred or performed self-narrative more successfully; revise her self-understanding; shift her self-narrative in deference to others; refute the perspectives of others by explaining away discrepancies with the projected self-narrative by pointing out interpretive errors on the part of others; or, finally, change the community for more accepting or tolerant for the projected self-understanding.

Although finding the equilibrium is crucial in Baylis' account of identity self-constitution, she emphasizes that fulfilling the equilibrium constraint by achieving stability of the self is not the end of "the game" of identity formation:

> To do so would be to contradict the claim that personal identity is a dynamic interpersonal communicative activity based on narrative and performance. Rather, equilibrium is a desired and desirable, temporary and temporizing, state of being that allows the self to take notice of her place in the world and through introspection and continued lived experience to refine or (radically) revise how she acts/performs in relations with others [64, pp. 123].

Thus, Baylis comprehends identity self-constitution as a process—she rejects a belief that there is the "true self" [3, 13, 17, 65–67] to be found; the core feature of relational identity is its dynamic nature. The self adopts more or less stable form depending on social, cultural, and political influences and pressures. What is important to emphasize is that over a person's lifetime "interludes of disequilibrium" are expected. Their intensity and frequency depends on: "(1) the person's skill in projecting who she is to herself as she lives her life and (more or less self–consciously) reflects on the actions and reactions that inform her interactions and transactions with others; and (2) the socio-cultural and socio-political context in which her actions, reactions, interactions, and transactions occur and are responded to" [64, pp. 123].

Up to the point discussed above, Baylis' account can be interpreted as an account that is primarily of a descriptive nature. The equilibrium constraint can be read not as the normative requirement for the narrative that is to be constitutive for personal identity, but as a theoretical postulate that allows to explain how the process of identity self-constitution works. Baylis [64], however, adds a normative flavor to her account by postulating that from an ethical perspective, it is crucial whether the state of equilibrium is achieved autonomously or as the result of oppression. For this reason, in neuroethical considerations about the potential effects of MMT on personal identity, the key issue in the light of Baylis' theory is whether a person undergoing memory modification can autonomously

constitute their identity or is subject to oppression. Such an approach is a traditional and, I contend, reliable guide to neuroethics.

In contrast, it seems that the accounts of Schechtman and Lindemann are problematic to apply in the neuroethical considerations. While I believe that both Schechtman and Lindemann accurately identify the relevant intuitions that usually govern the practices of self-constitution, and that it is often the case that people's judgements are implicitly driven by criteria postulated in their accounts, I nonetheless suggest that human self is too messy and dynamic to impose such demands on it. To elaborate on this, people have extraordinary abilities to fill in the memory and information gaps by producing narratives that are outrageously false but consistently (at least from their own perspective) explain why they are the way they are, did what they did, and how they found themselves in situations they found themselves in, as evidenced by decades of research on participants subjected to implementation of false memories [68, 69], patients who underwent commissurotomy (split brain patients) [70], or individuals suffering from Korsakoff's syndrome [71]. Moreover, research on self-deception shows that in typical population rationalizing and denying evidence that does not serve one's self is widespread [72]; even more, it is important adaptive [73] coping [74] mechanism, which, as some argue, is a consequence of the fact that deceit is fundamental in an animal communication [73]. Self-deception is arguably a product of the co-evolutionary struggle between deceiver and deceived selected for the evolutionary advantage of individuals who practice self-deception by allowing them to circumvent detection efforts [73]. It can therefore be concluded that cognitive mechanisms are conducive to preventing a typical person from meeting the reality constraint. For this reason, it does not seem adequate to a priori deny people authoritative self-constitution on the grounds of the reality constraint—acts of effective self-constitution could be too rare if it were to be so. Likewise, the credibility criteria postulated by Lindemann may not be appropriate to determine whether a given narrative has constitutive authority.

More generally, I suggest that any normative criterium postulated as required to constitute identity is susceptible to cultural biases. For example, both the reality and credibility constraints presuppose the significance of an ideal of leading "truthful" or "authentic" life—an ideal widespread in our culture but not necessarily culturally universal.[6] Thus, it may be that there is no normative condition that should be considered universally decisive for effective self-constitution. The process of self-constitution seems to be governed by different rule—that of "equilibrium" described by Baylis. Therefore, I will analyze the potential consequences of memory-modifying interventions to personal identity in the light of Baylis' account. It is worth emphasizing, however, that the intuitions regarding the practices of self-constitution identified by Schechtman and Lindemann are—as the following analysis will show—extremely useful for other purposes, i.e., predicting the reactions of others to changes in the narrative identity of a person who underwent an MMT intervention.

## The Ethics of MMTs: Autonomy and Baylis' Narrative Relational Account of Personal Identity

For Baylis, self-constitution requires minimal uptake of others, which "may be granted or withheld based upon judgments about 'reality,' 'coherence,' 'unity,' 'consistency,' 'credibility,' 'plausibility,' or the like, but what crucially matters given that we are constituted through our personal and public relations is the prospect of 'equilibrium'" [64, pp. 124]. Identity formation involves an iterative cycle of "self"-perception, "self"-projection, "other"-perception, and "other"-reaction. As I will show, memory modifications can have consequences at each of these stages. However, according to Baylis, this fact alone is not normatively problematic. For Baylis ethical issues arise only when given neurointervention affects autonomy. Thus, Baylis follows influential tradition according to which autonomy is central for the legitimization of practices in biomedical ethics.

The central idea behind the notion of autonomy is indicated by the etymology of the term: autós ("self") and nomos ("law" or "rule"). "Autonomy" was first applied to Greek city states exerting their own laws. This idea was later extended to person to describe

---

[6] This ideal may not be universal for some time even in WEIRD cultures, as postmodernists have long suggested that we live in an age of fluid identity; like politicians, we create our selves for the current needs of the 'social game,' and the truth as a guide in the process of self-creation has long lost its privileged status.

her (supposed) ability to "govern herself" or to self-determine, i.e., to decide and act (i.e., local autonomy), or more broadly to live (i.e., global autonomy), independent of influences that are not her own or are not expressive of her self [75]. Autonomy, then, is to be understood by specifying what the notion of "self" implies. Approaches to explaining the concept of "self" can, for the purposes of these considerations (see also [76, xi]), be divided into: (1) subjectivist or individualist accounts, highlighting subjective or individual traits and capacities of the person, (2) relational or interpersonal accounts, highlighting a person's social interdependencies, (3) narrative accounts, highlighting narrative practices as crucial in constructing and maintaining the self, and (4) embodied accounts, highlighting ways in which the body influences who we are and how we can be. Baylis' account has some elements of (1),[7] but it is primarily a combination of (2), (3) and (4) as she claims that: (i) "persons are (and can only be) dynamic complex co-creations informed by the perspectives and creative intentions of others" [64, pp. 118], (ii) one's personal identity "is a dynamic interpersonal communicative activity based on narrative and performance" [64, pp. 123], (iii) "race, class, gender, ethnicity, age, sexual orientation, and ability are features of the self that others read off the body in the context of 'complex networks of social norms, institutions, practices, conventions, expectations, and attitudes' [77, pp. 15]. So it is that 'agents' identities are formed within the context of social relationships and shaped by a complex of intersecting social determinants, such as race, class, gender and ethnicity' [78, pp. 4]"

Baylis [64, pp. 124] specifies that in relation to self-constitutive practices:

Autonomy is manifest when an individual actively contributes to authoring her life in a manner that is consistent with her broader interests, values, and commitments. How she lives in the world and succeeds in having others endorse or instantiate her self-projection is of pivotal importance. In addition, it matters for autonomy that an individual be able to decide which stories provided by others are to be incorporated into her self-narrative and whether to defend or revise her self-narrative in response to questioning, contesting, and possibly even refashioning of her identity by others.

By contrast, she claims that:

Oppression is manifest when an individual in certain contexts and circumstances is forced to live for periods of time within the confines of another's ideas about what makes for an appropriate self-narrative. An individual can be constrained by stories that are not of her own making but, rather, have been decreed, construed, and constructed by others who limit (in overt, covert, or insidious ways) who she can be, by actively and/or structurally restricting the range of narratives that can be appropriated and successfully enacted.

The above characterization of autonomy and oppression strongly resembles a constitutively relational analysis of Oshana [79, 80].[8] Therefore, it seems that Baylis' approach belongs to the category of strongly substantive relational accounts of autonomy. Such accounts are "constitutively relational" in a sense that social relations and exogenous circumstances constitute ineliminable elements of what self-determination involves. In other words, the state of the agent's external environment and interpersonal relationships are incorporated into the very definition of autonomy in such approaches [81].[9] Before considering whether and how memory modification can influence personal autonomy understood in this way, let me introduce the story of Tarana, which will serve to illustrate the potential ethical ramifications of using MMTs.

---

[7] This can be seen in the fact that Baylis claims that the intensity and frequency of disequilibrium partly depends on 'the person's skill in projecting who she is to herself as she lives her life,' the definition of autonomy below and the discussion in the section "MMTs and Autonomy Concern: Internal and External Autonomy."

[8] Many thanks to the reviewer for this suggestion.
[9] To see how this approach differs from proceduralist (a)historical and non-substantive analysis of autonomy, see [81].

## The Case of Tarana: Mapping the Potential Consequences of MMTs

Consider the story of Tarana[10] who was functioning seemingly well after experienced trauma, but as a result of experiencing intense stress in her thirties began developing symptoms of PTSD. Although Tarana suffered from sexual harassment in her childhood, she was able to build a meaningful life—she got involved into efforts to help girls living in marginalized communities. As Tarana acknowledges in her biography,[11] her traumatic experiences inspired her life-long pursuit to improve lives of young girls who undergo extreme hardships. The violent act she survived also motivated her to start the #MeToo movement. Therefore, the very aspect of her life that grew out of her traumatic memory was undoubtedly crucial to her autobiographical narrative, as she led her life story around this theme. Thus, her traumatic experiences constitute her self-defining memories. It is not unique since traumas are often critical in forming one's identity[12] [83, 84]. It should be noted that the crucial role of trauma in the process of identity formation is of great importance in the context of the ethics of therapeutic use of MMT, because in the most likely situations in which such interventions would take place, i.e. in the case of traumas, identity-related memories are the target of modification.

After Tarana had begun the #MeToo movement and shared her story during this process, she became haunted by vivid memories of the assault. This is all the more likely since survivors of sexual violence are experiencing a lot of stress after a disclosure, are often stigmatized, and verbally (or even physically) attacked. As Lindemann [85, pp. 42] puts it, the norm is the expectation that: "aspersions would be cast on her character, her motives, her credibility. […] even in the #MeToo era, the default assumption is that women's accusations of sexual assault are false, or that the assault wasn't that bad." Moreover, victims of sexual violence "do not benefit, to the same degree as other survivors, from telling their stories with the culturally valued narrative template of redemption" [85, 86]. Thus, as Tarana begun to be afraid that she will not be able to realize her mission to help others anymore—since her mental health started deteriorating due to intrusive and persistent recollections of traumatic memories—she decided to undergo treatment with the use of MMT. The potential therapeutic benefits and threats of Tarana's treatment would differ depending on the used MMT and memory system (and memory expression) that was targeted during the procedure. I will now consider two most relevant memory–modifying interventions in this context: modification of the memory valence and memory deactivation.

If MMT (e.g., propranolol) is used to alter Tarana's emotional responses towards traumatic memory, it could modify the valence of her traumatic memory and possibly relieve to some extent the symptoms of PTSD [87–90]. Such an intervention may even allow Tarana to unlearn her defensive responses because her emotional experience (and possibly physiological reactions) in response to cues associated with the traumatic event would be diminished [87–89, 91]. After this kind of intervention Tarana would, however, most likely retain conscious memory of the trauma along with the vivid details of this event (episodic memory) [29]. This assumption is additionally supported by evidence[13] provided by Brunet that PTSD patients undergoing propranolol memory modifications targeting emotional component of a trauma memory remembered their traumatic experience appropriately even when they had to recreate their trauma script from scratch (cited in [2]).

---

[10] Tarana Burke is the real name of the person who started the #MeToo movement. Although the above story is fictional, it was inspired by her life. The intention of choosing Tarana's story as an inspiration to these considerations about ethical ramifications of MMTs is to be able to put them in the context of real-life choices of the person involved in actions with significant ethical weight for a large number of people.

[11] Tarana Burke Biography (Accessed: 2021, December 19): https://www.biography.com/activist/tarana-burke

[12] One reason that it may be the case can be due to the side-effect of the cognitive adaptation according to which strongly aversive experiences are categorized as of greater significance. Since detecting and responding to threats is necessary for survival, excessively negative experiences constitute especially strong memory traces, i.e., they are remembered better than neutral experiences and are reactivated in a wider collection of situational contexts, thereby undergoing through the process of reconsolidation more often. For this reason, these memories strongly influence how one perceives and forms associations with environmental stimuli and, ultimately, how one's decisions are guided [82].

---

[13] Although this effect was observed during the experiment, it was not systematically studied.

On the other hand, if Tarana decides to undergo the intervention of selective erasure or deactivation[14] of the episodic memories of her traumatic experiences (e.g., with the use of optogenetic-like technology), such a memory-modifying procedure may prove unsuccessful in infringing the emotional component of these memories, as the emotionality of the memory seems to be consolidated in an independent non-declarative memory system in the form of associative memory with different neuronal underpinnings [92]. Thus, although Tarana could not consciously recollect her trauma as she previously did, she may remain with dormant disposition to experience emotional horror when the adequate cue is present. Finally, by erasing her episodic memories of trauma, Tarana may experience epistemic inaccessibility to the informational dimension of her self-defining memories.

Let me now identify potential implications of memory modifications for one's narrative identity. Assuming Tarana loses emotional component of her traumatic memories as a result of memory-modifying intervention—despite her knowledge of what was done to her—she could lose the ability to feel like a victim again. Consequently, she could not be able to self-identify with the story that have had at its center the idea that she experienced a trauma as she may feel that her identity is somehow "fake," since she could no longer relate to the crucial themes that built up her story on the emotional level. By contrast, by being deprived of the informational dimension of her self-defining memories due to erasure of episodic memories of trauma, Tarana may be left unable to properly articulate the autobiographical narrative that she had led throughout her life as she would lack the crucial information for constructing such a narrative—it's touchstone.

It could be argued, however, that this epistemic incompetence would not be a problem as long as Tarana would be able to learn of what was done to her through interaction with others as this would allow her to know her past. Furthermore, one could argue that when acquiring this information, she could relate to this truth on the emotional level—since the emotionality of her trauma would not be arguably infringed by the intervention—and thereby reconstruct her life story. On the other hand, reconnecting the emotional residues of Tarana's trauma to her life story might prove difficult, especially since others will hardly ever be able to give a full, extensive account of what occurred and how the individual experienced the situation.[15] Moreover, since the removed memories were episodic, even if it were possible for Tarana to acquire the truth about her past from a comprehensive account of what happened, she would presumably still lack at least some important dimensions of the episodic memory such as quasi-experientiality (that the memory includes spatiotemporal structure, perspectivity, and modality-specific sensory information) and autonoetic consciousness (that the event information is meta-represented as having been obtained firsthand) [93], which accompanied her dwelling on the events of mistreatment. Therefore, even if the information about her traumatic past were disclosed to her by third parties—which itself is controversial from an ethical perspective as Tarana underwent memory-modifying interventions precisely in order to forget about it—she could still be deprived of the ability to believe that she, herself, experienced the trauma, as the important features of the phenomenology usually accompanying memories that inform the person about encountering events first-handedly would be inaccessible to her. Thus, similar to the modification of the emotionality of the memory, memory erasure can lead to the inability to self-identity with the life story that has been in force so far.

The first thing to notice in this context is that the sole fact that Tarana may feel or hold beliefs at odds with the story that had been prevailing in her autobiographical narrative prior to memory modification is not problematic by the lights of Baylis' approach. This is because Baylis endorses the idea that maintaining personal identity is a dynamic process—self-narratives can change and it is expected that the person experiences interludes of self–disruption. More important is how the person responds when her operative autobiographical story starts to get problematic.

One way Tarana could respond to the narrative disruption she experiences as a result of memory modification intervention is to keep the foregoing narrative. In the case of the valence modification, the episodic as well as semantic memory of the experienced

---

[14] Deactivation, meaning that she can no longer retrieve them, although they may be recoverable through repeated intervention as, e.g., optogenetics effects may be, in principle, reversed.

[15] Many thanks to the reviewer for suggesting this argument.

trauma would remain intact; thus, maintaining the previous narrative appears as an option in such a situation since Tarana would have all the necessary information to do so. However, given that the success of Tarana's therapy relies on her losing the accessibility to the emotional dimension of traumatic memories that were defining to her self-understanding, although she may be able to articulate her previous story, she may be unable to emotionally relate to her narrative in the aftermath of the intervention. Thus, it is plausible that she would decide to give up the past narrative. Similarly, in the case of the erasure of traumatic memories, as a result of not being able to believe or even articulate her previous life narrative due to insufficient information, Tarana may not be able to lead her previous autobiographical story. It is also worth emphasizing that if there were the possibility of a combined intervention that would both remove the informational content and the emotional residues of the memory, abandoning the former narrative identity would be even more likely in this case.

Thus, stripped of the previous narrative, Tarana would face the dilemma of a person without a life story. So what could she do? Given that making sense of memories and experiences in narrative terms seems to be a fundamental human need, she would arguably construct a novel autobiographical narrative. But what narrative can the person create whose life so far has revolved around a memory that is no longer there? She could, for example, adopt narrative identity that is consistent with one of the prevalent ideologies of "good life" of our times, such as that promoted by late capitalism (Lindemann's "the given"). As a result, Tarana could be prompt to give up the work in the underprivileged communities in order to follow a more profitable career. Now she uses legal skills she acquired during her education and a career as an activist to work for corporations by helping them to escape to tax havens (financially viable profession). From now on Tarana perceives herself as a corporate lawyer and tell a story about herself in which she has a profit-oriented ideas on how to best live one's life. This may seem as a quite radical change for Tarana. One may argue that motivational foundation of Tarana's activism was rather shallow if such serious change were possible—one could expect that Tarana's sense of justice, empathy for other victims, or some other robust motive to the cause of ending sexual violence against women should move her towards her altruistic pursuits, and not just her personal memories. However, taking into account that the most passionate and persistent people who fight for social change are those who themselves have experienced persecution, discrimination, etc., it may just be a psychological fact that a strong first-personal emotional experiences move persons to actions and long-term commitment to the cause.[16] Moreover, as mentioned above, such emotional (often traumatic) memories—the very memories that were targeted during the procedures—are often at the very heart of one's identity. Therefore, a global narrative shift seems likely as a result of modulating such memories.

The above interpretation of Tarana's case is an attempt to map the potential consequences of memory modifications by showing that they can affect various stages of identity formation identified by Baylis (i.e., "self"-perception, "self"-projection, "other"-perception, "other"-reaction). However, as mentioned above, establishing this does not prove that MMTs are ethically problematic. According to the Baylis' model, only identity-related changes that affect a person's autonomy are relevant to the ethical evaluation of such interventions. Thus, I will now discuss how memory modification can influence personal autonomy.

## MMTs as a Means of Increasing Autonomy

Tarana's case shows that memory modification can, in some significant ways, positively affect global autonomy, namely, strengthen it. Memory-modifying interventions served Tarana as means to overcome her feeling (or belief) of being constrained by traumatic experiences that unwillingly fixed her ideas about who she is and who she can be.[17] Exercising autonomy in terms proposed by Baylis presupposes

---

[16] See Lavazza [94] for an in-depth discussion of this issue in the context of the ethics of memory modification.

[17] Obviously, as discussed above, the fact that some memories continue to have such an effect on Tarana may not only be due to internal factors of her psychological constitution, but also to external factors—"the given," i.e., that they may not be adequate and commonly accepted narrative templates in our culture for victims of sexual violence that would help them to work through their traumas.

that a person has the ability to critically reflect on self-narrative she internalized, as well as reject it.[18] In the case of Tarana, due to intrusive memories, she did not possess the latter of these crucial capacities before memory modification. She had not been able to reject internalized self-narrative. This interpretation is likely since Tarana suffered from PTSD, which is often categorized as a paradigmatic example of hijacking narrative identity [95, 96]. In extreme cases, individuals with PTSD can act and feel as if they are trapped in the traumatic situation [2].[19] By easing or erasing memories MMTs could help a person overcome their maladaptive narrative scripts that kept them from letting things go and more autonomously authoring their lives in a manner that they find more consistent with their broader interests, goals and values.

As discussed in the previous section, Tarana underwent an MMT intervention to be able to continue helping others. Imagine that this scenario was a bit different, namely that Tarana's motivation to modify her memory was due to more egoistic rather than altruistic impulses. Despite all her work as an activist, Tarana could feel that she was in that role only because of the paralyzing memories that allowed her to do nothing else, but built an entire identity around this theme. Soothing impact of these memories were to allow her to change this oppressive, internalized self-understanding. And indeed, memory modification could help Tarana to achieve this goal by changing her "self"-perception and provide capacities for a novel "self"-projection. Devoid of traumatic memories, she may stop perceiving herself as a victim and feel limited from following a radically different life-path such as becoming a corporate lawyer.

This notion can be further specified with Frankfurt's hierarchical model of desires and volitions [97, 98]. According to Frankfurt's account, autonomy requires that a person is capable to reflect and approve (or disapprove) her first–order desires in the

light of her second–order desires and to form volitions—and consequently actions—on the basis of endorsed desires. Although Tarana was able to reflect critically on her first–order desires that were related to her identity as a social activist (e.g., helping other victims, fighting with perpetrators etc.), she was not able to act in line with her second–order desires, i.e., following a different life-path. This is because her lower desires subdued the implementation of her self-creation projects, as she was constantly exposed to the unwanted evocation of traumatic memories in her daily life. Although this was at the expense of her own global autonomy, she could not help but engage in helping other victims. Maslen et al. [99] discuss a strategy that could help with this kind of issues related to autonomy in the context of deep brain stimulation (DBS) therapy for anorexia nervosa. They propose that DBS can serve as a means to reduce symptoms of anorexia by "promoting comparative cognitive control." Comparative cognitive control in anorexia can be enhanced with DBS either by reducing the compulsive need to diet, or by increasing top–down control over compulsion to avoid eating. In the case of memory modification, the former kind of these strategies seems applicable. By targeting the source of Tarana's first-order desires, i.e., her traumatic memories, memory modification can reduce (or even eliminate) the urge to yield to the desires that arise from such memories, and, as a result, the need of building the self-narrative around them. In this way, memory modification could help Tarana to regain control over forming volitions and acting in accordance with her higher–order desires. Consequently, a memory modification intervention could enable her to construct a novel self-understanding, which would enhance her autonomy in shaping her identity.

## MMTs and Autonomy Concern: Social Aspects

However, there will arguably be a tension between Tarana's preferred self-narrative and the performed self-narrative, that is, the story of who she can be "given the ways in which her life is constrained by self and others" [64, pp. 119]. It is because, it is most likely that the uptake of her community ("other"-reaction) will be almost unambiguously negative since values and life-project (being a corporate lawyer) around which Tarana is building her life and narrative

---

[18] This is the subjectivist component of Baylis account.

[19] As the reviewer rightly pointed out: Since the narrative encompasses not just what happened in a person life but also the interpretation, meaning and relevance of those events for the person, there seems to be a room to reject the narrative on those levels. While in normal circumstances a person may be able to reject the interpretation and self-defining power of the event, in the case of traumatic memories, especially those underlying PTSD, the person may not possess this capacity.

after memory modification stand in stark opposition to those that built her previous identity (a social activist). This claim is supported by the X-Phi results. As Dranseika [64, 100] neatly puts it:

Empirical research on folk reasoning about authenticity finds again and again that moral properties are at the very core of people's judgments of who the person is "deep down inside" and whether they are "still the same person" after undergoing some transformation. Interestingly, depending on what moral beliefs the person issuing a judgment holds, the same scenario (e.g. a religious person becoming an atheist, unpatriotic person becoming patriotic) can be perceived as either a situation in which the true self is expressed or suppressed. Furthermore, when a change of moral character occurs, the direction of change is crucial: in cases of transformations involving moral improvement, study participants are more inclined to judge the post transformation individual to be still the same person compared to cases involving moral deterioration.

Tarana's case is representative of all these features: her transformation includes moral properties, her community holds moral beliefs that are in opposition to moral beliefs she acquires as a result of memory modification, and the direction of change involves moral deterioration (at least, in the eye of the beholder—her community). Moreover, taking into account the above considerations that the approval of others is granted on the basis of judgments about such characteristics of self-narrative as, e.g., coherence, unity, consistency, and credibility, makes a positive reception of Tarana's self-understanding even less likely as the extent and abruptness of change in Tarana's post-interventional life story largely questions these features.

Moreover, it matters whether Tarana would be able to motivate appropriate uptake of the self-narrative that she values after memory modification in any community. The problem with the abandonment of one's community and entering the new one is, for example, that in most communities a long-term bond and shared years of experience are crucial for the full acceptance of the member. For this reason, changing the community to which one belongs is difficult, stressful, and not always possible, therefore many concessions can be expected from individuals, i.e.,

sacrificing a certain degree of autonomy in shaping their identity just to stay in-group.

It seems that the notion of 'community' as used by Bailys and in the above considerations implies a certain degree of unity and similarity among its members. In the case of Tarana, it appears plausible that some members of her community would tolerate or endorse her new narrative or that she would include some people into her community that would.[20] Accordingly, the account of Baylis should be expanded to include a definition of what and who is (part of) the community with which to find the equilibrium. I suggest that people with whom individuals (should) care to find the equilibrium are *significant others*, i.e., others who are considered to have special authority, within a particular group or context, with regard to co-creating their identity. This authority consists in the power to legitimately endorse or challenge (i.e., dismiss, resist, contest) (parts of) one's self-narrative. According to Westlund [101], one of the necessary conditions for legitimate challenge is its relational situatedness: a legitimate challenge must be situated in sense-giving relationship: "some sense-giving relationships are broad: one is a member of the moral community, citizen of the nation, inhabitant of a community. Others are more narrow: one is a mother, husband, neighbor, or club member. But at very least, it must be clear why it *matters* to my critic why I think and act the way I do, and it must matter to her in a way that she can reasonably expect to matter to me." To illustrate these theoretical considerations with an example: it is doubtful that colleagues are in a position to legitimately challenge (i.e., dismiss, resist, contest) one's narrative about being a good daughter, and vice versa, during a family dinner, the parental praise (i.e., active endorsement) of a child's scientific achievements will not add to the narrative that the child is a great philosopher who surely deserves an award from the philosophy faculty (unless, perhaps, the parents are deans of the faculty).

One may argue that the advent of social media has complicated this picture as, e.g., what used to be the most private and intimate is now often the subject of fierce public and political debate. This issue

---

[20] Thanks to the reviewer for this insight.

is particularly evident in the case of a public figure like Tarana, where many people (even strangers) could arguably reject and challenge or embrace her self-narrative.[21] One of the important questions that arises in this context is whether achieving equilibrium on social media is sufficient for establishing identity-constituting narrative. It seems that Baylis' account leaves this possibility open, which may be a positive result in view of the fact that social media identity discussions are often emotional, suggesting that people care about them. My suggestion on this topic is that in fact it is an open matter which will depend on whether significant others participate in such a discussion.

The above considerations suggest that although our identity may appear more and more globalized with the age of social media, intimate, sense-giving relationships are still essential for constituting identity. If we look at how people establish and maintain such relationships, these processes are governed by certain psychological regularities, and narratives about past are in its heart. Imagine a person with whom you have a great time, but whose life you know nothing about—hardly anyone would call such a person a partner or friend. People usually call a partner or friend the person in whose life they participate, so the partner or friend's life story unfolds before their eyes, or at least they need to hear stories or relevant parts of stories they do not know to call the person a partner or friend. So it can be said that the narratives of our past experiences are what binds love and friendship. Taking this into account, the closer we get to someone, the more destructive memory modification can be for our relationship. This, on the other hand, may have consequences not only for a person's happiness but also for her autonomy. A person with fewer sense-giving relationships—whether they are broad, but this is especially true of the narrow ones, i.e., a partner or friends—has less support and a weaker social position,[22] which reduce her autonomy in shaping her life and identity.

## MMTs and Autonomy Concern: Internal and External Autonomy

The above considerations suggest that there are at least two important senses of exercising autonomy that one should distinguish when considering the ethics of memory modification: that a person has sufficient reflective capacities and emotional predispositions to create and sustain preferred self-narrative and negotiate a story that is to be incorporated into her performed self-narrative (*internal identity-related autonomy*), and that a person lives in sufficiently open society (not necessarily only communities) to be able to carry out her preferred self-narrative in a way that does not enforce her to live within the confines of other's ideas (*external identity-related autonomy*). What is relevant, substantive relational accounts of autonomy, such as Baylis' approach, have no problem incorporating such a dual-basis framework. Accounts of this sort "agree that portraits of personal autonomy must include reference to the structure of the agent's psychology, the agent's competencies, and the nature and origin of the psychological states of the agent" as well as "mandates that certain substantive social arrangements and social roles be present, and relations of subordination and subservience, and economic or psychological insecurity, must be absent" [81].

Although by modifying her memory Tarana may be able to gain emotional or cognitive predispositions allowing her to free herself from one kind of oppression, i.e., internal oppression that her memories exerted on her, she may not be able to free herself from oppression (or, to put it mildly, social pressure) that comes from the society. In other words, although memory modification may influence the agent's mental life in a way that makes a person more autonomous, i.e., strengthen her internal autonomy, it cannot change others to be more accepting, i.e., enhance external autonomy. By the lights of Baylis' account this is, however, important consideration, since in order to be able to achieve equilibrium, one's project of self-creation must be endorsed by others. In other words, "internal emigration" in search of freedom is not enough as it does not allow to effectively constitute one's identity. This fundamental interdependence regarding self-constitution shows that social factors can often be decisive in assessing whether memory-modifying procedure is ethically desirable.

---

[21] Many thanks to the reviewer for the suggestion to clarify the concept of 'community' and this observation about Tarana, which prompted me to extend Baylis' account.

[22] DeScioli and Kurzban [102] argue that friendship evolved precisely to receive support in potential conflicts.

In this context, it is worth noting that, on the face value, a person with increased internal autonomy should be in a better position to exercise their external autonomy. As Baylis notes, identity formation is an iterative and cyclical process in which a person unsatisfied with the uptake of her self-narrative can take proactive measures. Since Tarana's projected self-narrative may be more or less well-read depending on, e.g., the access of third parties to interpretative facts about Tarana, she may try to project her preferred self-narrative in a way that shows that her narrative change is a coherent part of an ongoing life story, and memory modification was only a tool that allowed her for the projection of this narrative.

However, taking into account that third-parties build a picture of a person basing only on certain manifestations or aspects of one's self (behavior, performed narrative etc.), a lot is hidden from them—people usually know much more about their own mental states (e.g., desires, memories, traumas, future plans, dreams) than about mental states of others. Thus, since others base their image ("other"-perception) of a person only on "observational facts"—the ones that were revealed in the behavior of a person over the course of years—they may have a tendency to be more conservative regarding the endorsement of other's self-constitutive projects, which depict radically different picture of that person (especially when this change includes features emphasized by Dranseika [100], or is in tension with such indicators as "unity," "reality," "credibility," etc.). Therefore, taking into account that memory modification interventions may often concern self-defining memories, that is, memories largely responsible for the shape of a person's life—as in the case of Tarana—sudden narrative and behavioral changes that can be the result of such procedures may be hard to explain to significant others. For this reason, a person after memory-modifying procedure who experienced radical transformation may encounter a lot of resistance from her community to endorse her novel life-story ("other"-reaction). To put it briefly, improbable testimonies on behalf of self-understanding must be especially good if they are to be accepted. But can a person, after memory modification, hope to achieve this—that is, present a compelling story of her life?

It seems problematic because a person after memory modification may lose a significant part of self-knowledge, and as a result knows less about her self-defining memories and previous life-story than significant others. The story of Tarana, who, in an attempt to make her story more convincing for third parties, depicts herself as "a person who fights with the demons of the past and at the moment she feels she may lose this battle, she decides for the treatment that gives her best chances to let the foretime go" could be regarded as particularly convincing testimony as it emphasize coherence, unity, consistency of Tarana's life-story and, at the same time, expose potential interpretative mistakes on the third-parties end. However, the very skill to project such self-understanding relies on Tarana being able (or willing) to build her narrative based on the very memories that were erased (or modified) during the procedure. It seems that in the case of valence modification, Tarana may be able to present such a story. However, taking into account that she probably would not identify with it, her motivation to do so is questionable. In the event that the memories are erased, she may not be able to articulate such a testimony at all.

Moreover, when it comes to "other"-reaction, there is a difference for third-parties between accepting that Tarana no longer wants to build her identity in relation to the negative experiences that have happened to her and a positive uptake (e.g., endorsing, supporting) of her new identity (a corporate lawyer) that goes against their values. In the case of Tarana, the most likely solution on behalf of significant others would be to invite her to revise her projected self-narrative. After all, they "owe" her great debt of gratitude and it is unlikely that they would immediately want to banish her from the community (thereby Tarana would arguably be in a more privileged position in this respect than an average person). However, as mentioned above, as a result of memory modification Tarana may either lose the emotional component that determines the very motivation to engage in the identity game with the people who were significant others in her life (she may want to find more suitable community), or she may lose the tools that would allow her to effectively participate in it. If added to this that third-parties in such a situation has the informational advantage over Tarana in terms of knowing her previous self-defining memories (traumas), and that they may have a vested interest in keeping her in the community in her previous role (as she is the community founding member), by dismissing, resisting, or actively contesting her novel narrative

Tarana's social surroundings could have a disproportionate large influence on the shape and the content of self-narrative she perceives as possible to perform. In other words, memories can be seen as the currency that enables individuals to be both autonomous and successful in the identity game. By depriving themselves of this currency by a memory modification intervention, individuals run the risk of being in a worse epistemic and negotiating position. It is easier to influence and rule someone who does not have a complete picture of one's life. Taking this into consideration, it seems doubtful to assume that memory modulation would enhance one's external autonomy.

I am far from saying that the question of the openness of significant others or, more broadly, the society in which one lives determines the ethical evaluation of a given memory-modifying intervention. However, what Baylis' account suggests is that the possibility of autonomous self-constitution greatly depends on this social factor. Therefore, it must be taken into account that greater inner freedom (internal autonomy) does not have to translate into greater overall autonomy. By contrast, with greater inner freedom can, in some circumstances, come greater oppression from the outside—e.g., when the projected narrative of a person after a memory-modifying intervention is in opposition to the values that that person professed before the intervention and that are still hold by the significant others. Thus, we should identify and inform the patients about the dangers of memory-modifying interventions, such as the possibility of experiencing social pressure, stigmatization or even oppression, which can sometimes make it impossible for the patient to shape and achieve the desired identity.

## Conclusions

This article has examined the question of whether and how memory modifications, which may soon become reality due to the memory-modifying prospects promised by the development of MMTs, may threaten personal identity and autonomy. Both modification of the valence and memory deactivation have been shown to have the potential to affect each stage of the identity formation cycle ("self"-perception, "self"-projection, "other"-perception and "other"-reaction). Each of these stages can be influenced by the memory modifications to such an extent that such interventions can

hinder a person's efforts to autonomously structure and constitute her identity. I distinguished between internal and external forms of identity-related autonomy and oppression, and argued that while memory modification may, in certain contexts, be helpful in combating internal oppression, i.e., constraints exerted on the person's psychology by her (traumatic) memories, it can often have negative consequences for external autonomy—modifying a person's memory will not change the social context in which she lives, and others reception is crucial in any relational model of personal identity.

This article provides the tools to identify the challenges and risks for when the person returns to the community after memory-modifying interventions; it shows what social circumstances may particularly threaten the autonomy in shaping and maintaining personal identity. Taking into account that social groups (e.g., families, friends, workplace, academia) constitute their very existence around given values, goals, taboos, etc., they necessarily limit acceptable self-narratives of their members. It means that a person whose memories were erased or modified and whose self-understanding radically changed may not be able to obtain minimal uptake by significant others and find equilibrium in a given community. This can lead an individual to "find herself forced to live within the confines of another's ideas about what makes for an appropriate self-narrative" [64, pp. 124], or abandon her social group, which can also limit a person's overall autonomy. Since "one's position in society, in turn, affects who one can be" [64, pp. 124], by leaving one's community, a person condemns herself to be a newbie in a new social surroundings, which can limit her negotiating position and the prospect of achieving satisfactory equilibrium.

Moreover, even in the case when a person declares, prior to undergoing memory-modifying intervention, that she prefers to be free from internal oppression of memories than to stay in a given group, the wider social context still matters, that is, it matters whether there are significant others within her groups or other groups valuable from the person's perspective that will accept her new identity, disregarding her past. Although in a liberal society there is an extensive market of social groups from which one can choose, relevant factors that made the members of the group strengthen the bond may be unavailable to the newbie (e.g., shared experience, duration of the relationship,
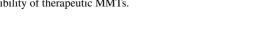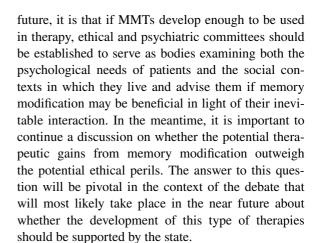
trust). Moreover, some groups (e.g., family, friends) and significant others (e.g., mother, partner) may not be replaceable (psychologically and otherwise).

Memory modifications run the risk that a person will not be able to achieve the equilibrium, and thus meet one of the most fundamental human needs—the autonomous construction and maintenance of personal identity. However, if we take moral principles of political liberalism seriously, this is not sufficient to justify forbidding citizens from undergoing memory modifications. If such interventions are possible, the right to control one's own mental life, and thus the decision to undergo medical intervention that interferes with one's own memory, should be in the hands of the patient. In other words, when making decisions about memory modification, the patient's local autonomy overrides potential concerns about ramifications for her global autonomy. This assessment could change if such interventions would significantly increase the risk of harm to others [1]. However, this article does not discuss the type and degree of harm to others that would be sufficient to prevent persons from using MMTs. This consideration leads to the question of whether: "does it follow, from the fact that memory erasure will usually be permissible, that moral assessment of it is pointless?" As Levy [16] reminds us: "no: we assess things from the moral point of view for many reasons, not just to see whether they should be banned. There are, after all, actions intermediate between forbidding and ignoring: we can advise, exhort, praise and disapprove; with regard to the behaviour of others and of ourselves."

To conclude, although the state should not be entitled to prohibit the use of MMTs—especially in a therapeutic context,[23] that is, in light of the fact that the identity of the potential patient likely to be affected by the intervention is not healthy (or even functional) in such a context—the potential patient should be thoroughly informed of the dangers of such interventions, because the changes she needs to be prepared for are extensive. As discussed in this paper, these changes may include the resignation from relationships that have been crucial for a person and the need to restructure one's life and identity. If these considerations can teach us any lessons for the future, it is that if MMTs develop enough to be used in therapy, ethical and psychiatric committees should be established to serve as bodies examining both the psychological needs of patients and the social contexts in which they live and advise them if memory modification may be beneficial in light of their inevitable interaction. In the meantime, it is important to continue a discussion on whether the potential therapeutic gains from memory modification outweigh the potential ethical perils. The answer to this question will be pivotal in the context of the debate that will most likely take place in the near future about whether the development of this type of therapies should be supported by the state.

**Declarations**

**Conflict of Interest**  I have no conflict of interest to declare.

---

[23]  But see [24] for a more nuanced discussion of the permissibility of therapeutic MMTs.

### References

1. Adamczyk, Agnieszka K., and Przemysław Zawadzki. 2020. The Memory-Modifying Potential of Optogenetics and the Need for Neuroethics. *NanoEthics* 14: 207–225. https://doi.org/10.1007/s11569-020-00377-1.
2. Elsey, James, and Merel Kindt. 2016. Manipulating Human Memory Through Reconsolidation: Ethical Implications of a New Therapeutic Approach. *AJOB*

*Neuroscience* 7: 225–236. https://doi.org/10.1080/21507740.2016.1218377.

3. Erler, Alexandre. 2011. Does Memory Modification Threaten Our Authenticity? *Neuroethics* 4: 235–249. https://doi.org/10.1007/s12152-010-9090-4.

4. Farah, Martha J., Judy Illes, Robert Cook-Deegan, Howard Gardner, Eric Kandel, Patricia King, Eric Parens, Barbara Sahakian, and Paul Root Wolpe. 2004. Neurocognitive enhancement: What can we do and what should we do? *Nature Reviews Neuroscience* 5: 421–425. https://doi.org/10.1038/nrn1390.

5. Gligorov, Nada. 2016. *Neuroethics and the scientific revision of common sense*. New York: Springer.

6. Henry, Michael, Jennifer R. Fishman, and Stuart J. Youngner. 2007. Propranolol and the Prevention of Post-Traumatic Stress Disorder: Is it Wrong to Erase the "Sting" of Bad Memories? *The American Journal of Bioethics* 7: 12–20. https://doi.org/10.1080/15265160701518474.

7. Hui, Katrina, and Carl E. Fisher. 2015. The ethics of molecular memory modification. *Journal of Medical Ethics* 41: 515–520. https://doi.org/10.1136/medethics-2013-101891.

8. Iftode, Cristian, Alexandra Zorila, Alexandra, Constantin Vica, and Emilian Mihailov. 2022. Experimental and Relational Authenticity: How Neurotechnologies Impact Narrative Identities. *Phenomenology and the Cognitive Sciences*: 1–18. https://doi.org/10.1007/s11097-022-09825-7.

9. Kass, Leon. 2003. *Beyond therapy: Biotechnology and the pursuit of human improvement*. The President's Council on Bioethics.

10. Kolber, Adam J. 2006. Therapeutic Forgetting: The Legal and Ethical Implications of Memory Dampening. *Vanderbilt Law Review* 59: 1561–1626.

11. Kroes, Marijn C. W., and Rain Liivoja. 2019. Eradicating war memories: Neuroscientific reality and ethical concerns. *International Review of the Red Cross* 101: 69–95. https://doi.org/10.1017/S1816383118000437.

12. Lavazza, Andrea. 2018. Memory-modulation: Self-improvement or self-depletion? *Frontiers in Psychology* 9: 469. https://doi.org/10.3389/fpsyg.2018.00469.

13. Lavazza, Andrea. 2019. Moral Bioenhancement Through Memory-editing: A Risk for Identity and Authenticity? *Topoi* 38: 15–27. https://doi.org/10.1007/s11245-017-9465-9.

14. Leuenberger, Muriel. 2022. Memory Modification and Authenticity: A Narrative Approach. *Neuroethics* 15: 10. https://doi.org/10.1007/s12152-022-09489-9.

15. Levy, Neil. 2007. *Neuroethics: Challenges for the 21st Century*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511811890.

16. Levy, Neil. 2011. Changing One's Mind The Ethics of Memory Erasure in Eternal Sunshine of the Spotless Mind. *S&F_Scienzaefilosofia.it* 5: 20–49.

17. Liao, S. Matthew., and Anders Sandberg. 2008. The Normativity of Memory Modification. *Neuroethics* 1: 85–99. https://doi.org/10.1007/s12152-008-9009-5.

18. Liao, S. Matthew., and David T. Wasserman. 2007. Neuroethical Concerns about Moderating Traumatic Memories. *The American Journal of Bioethics* 7: 38–40. https://doi.org/10.1080/15265160701518623.

19. Parens, Erik. 2010. The Ethics of Memory Blunting: Some Initial Thoughts. *Frontiers in Behavioral Neuroscience* 4: 190. https://doi.org/10.3389/fnbeh.2010.00190.

20. Wasserman, David. 2004. Making Memory Lose Its Sting. *Philosophy and Public Policy Quarterly* 24: 12–18. https://doi.org/10.13021/G8pppq.242004.196.

21. Wolpe, Paul R. 2002. Treatment, enhancement, and the ethics of neurotherapeutics. *Brain and Cognition* 50: 387–395. https://doi.org/10.1016/S0278-2626(02)00534-1.

22. Vukov, Joseph. 2017. Enduring Questions and the Ethics of Memory Blunting. *Journal of the American Philosophical Association* 3(2): 227–246. https://doi.org/10.1017/apa.2017.23.

23. Zawadzki, Przemysław, and Agnieszka K. Adamczyk. 2021. Personality and authenticity in light of the memory-modifying potential of optogenetics. *AJOB Neuroscience* 12: 3–21. https://doi.org/10.1080/21507740.2020.1866097.

24. Zawadzki, Przemysław, and Agnieszka K. Adamczyk. 2021. To remember, or not to remember? Potential impact of memory modification on narrative identity, personal agency, mental health, and well-being. *Bioethics* 35: 891–899. https://doi.org/10.1111/bioe.12926.

25. Kindt, Merel. 2014. A behavioural neuroscience perspective on the aetiology and treatment of anxiety disorders. *Behaviour Research and Therapy* 62: 24–36. https://doi.org/10.1016/j.brat.2014.08.012.

26. Schiller, Daniela, Marie-H. Monfils, Candace M. Raio, David C. Johnson, Joseph E. LeDoux, and Elizabeth A. Phelps. 2010. Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature* 463: 49–53. https://doi.org/10.1038/nature08637.

27. Yehuda, Rachel, Linda M. Bierer, Laura C. Pratchett, Amy Lehrner, Erin C. Koch, Jaklyn A. Van Manen, Janine D. Flory, Iouri Makotkine, and Tom Hildebrandt. 2015. Cortisol augmentation of a psychological treatment for warfighters with posttraumatic stress disorder: Randomized trial showing improved treatment retention and outcome. *Psychoneuroendocrinology* 51: 589–597. https://doi.org/10.1016/j.psyneuen.2014.08.004.

28. Soravia, Leila M., Markus Heinrichs, Livia Winzeler, Melanie Fisler, Wolfgang Schmitt, Helge Horn, Thomas Dierks, Werner Strik, Stefan G. Hofmann, and Dominique J.-F.. de Quervain. 2014. Glucocorticoids enhance in vivo exposure-based therapy of spider phobia. *Depression and Anxiety* 31: 429–435. https://doi.org/10.1002/da.22219.

29. Phelps, Elizabeth A., and Stefan G. Hofmann. 2019. Memory editing from science fiction to clinical practice. *Nature* 572: 43–50. https://doi.org/10.1038/s41586-019-1433-7.

30. Giustino, Thomas F., Paul J. Fitzgerald, and Stephen Maren. 2016. Revisiting propranolol and PTSD: Memory erasure or extinction enhancement? *Neurobiology of learning and memory* 130: 26–33. https://doi.org/10.1016/j.nlm.2016.01.009.

31. Alberini, Cristina M., and Joseph E. LeDoux. 2013. Memory reconsolidation. *Current Biology* 23(17): R746–R750. https://doi.org/10.1016/j.cub.2013.06.046.

32. Ehlers, Anke, and David M. Clark. 2000. A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy* 38: 319–345. https://doi.org/10.1016/s0005-7967(99)00123-0.

33. Deisseroth, Karl. 2012. Optogenetics and psychiatry: Applications, challenges, and opportunities. *Biological Psychiatry* 71: 1030–1032. https://doi.org/10.1016/j.biopsych.2011.12.021.

34. Hernandez, Pepe J., and Ted Abel. 2008. The role of protein synthesis in memory consolidation: Progress amid decades of debate. *Neurobiology of learning and memory* 89: 293–311. https://doi.org/10.1016/j.nlm.2007.09.010.

35. Josselyn, Sheena A., and Susumu Tonegawa. 2020. Memory engrams: Recalling the past and imagining the future. *Science* 367. American Association for the Advancement of Science. https://doi.org/10.1126/science.aaw4325.

36. Ramirez, Steve, Xu Liu, Pei-Ann. Lin, Junghyup Suh, Michele Pignatelli, Rorger L. Redondo, Tomás J. Ryan, and Susumu Tonegawa. 2013. Creating a False Memory in the Hippocampus. *Science* 341: 387–391. https://doi.org/10.1126/science.1239073.

37. Redondo, Roger L., Joshua Kim, Autumn L. Arons, Steve Ramirez, Xu Liu, and Susumu Tonegawa. 2014. Bidirectional switch of the valence associated with a hippocampal contextual memory engram. *Nature* 513: 426–430. https://doi.org/10.1038/nature13725.

38. Chen, Shuo, Adam Z. Weitemier, Xiao Zeng, Linmeng He, Xiyu Wang, Yanqiu Tao, Arthur J. Y. Huang, et al. 2018. Near-infrared deep brain stimulation via upconversion nanoparticle-mediated optogenetics. *Science* 359: 679–684. https://doi.org/10.1126/science.aaq1144.

39. Wang, Shutao, Tara Kugelman, Amanda Buch, Mathieu Herman, Yang Han, Maria Eleni Karakatsani, S. Abid Hussaini, Karen Duff, and Elisa E. Konofagou. 2017. Non-invasive, focused ultrasound-facilitated gene delivery for optogenetics. *Scientific Reports* 7: 39955. https://doi.org/10.1038/srep39955.

40. Bedbrook, Claire N., Kevin K. Yang, J. Elliott Robinson, Elisha D. Mackey, Viviana Gradinaru, and Frances H. Arnold. 2019. Machine learning-guided channelrhodopsin engineering enables minimally invasive optogenetics. *Nature Methods* 16: 1176–1184. https://doi.org/10.1038/s41592-019-0583-8.

41. Rich, Megan, Eric Zhang, Ashley Dickey, Haley Jones, Kelli Cannon, Yuriy Bandera, Stephen Foulger, Farah Lubin, and Mark Bolding. 2020. A noninvasive approach to optogenetics using focused ultrasound blood brain barrier disruption for the delivery of Rradioluminescent particles. *bioRxiv*: 2020.08.20.248302. https://doi.org/10.1101/2020.08.20.248302.

42. Gong, Xin, Diego Mendoza-Halliday, Jonathan T. Ting, Tobias Kaiser, Xuyun Sun, André M. Bastos, Ralf D. Wimmer, et al. 2020. An ultra-sensitive step-function opsin for minimally invasive optogenetic stimulation in mice and macaques. *Neuron* 107: 38-51.e8. https://doi.org/10.1016/j.neuron.2020.03.032.

43. Sahel, José-Alain, Elise Boulanger-Scemama, Chloé Pagot, Angelo Arleo, Francesco Galluppi, Joseph N. Martel, Simona Degli Esposti, et al. 2021. Partial recovery of visual function in a blind patient after optogenetic therapy. *Nature Medicine* 27: 1–7. https://doi.org/10.1038/s41591-021-01351-4.

44. Adamczyk, Agnieszka K., and Przemysław Zawadzki. 2021. Personality and authenticity in light of the memory-modifying potential of optogenetics: A reply to objections about potential therapeutic applicability of optogenetics. *AJOB Neuroscience* 0: 1–4. https://doi.org/10.1080/21507740.2021.1938292.

45. Gilbert, Frederic, and Eliza Goddard. 2014. Thinking ahead too much: Speculative ethics and implantable brain devices. *AJOB Neuroscience* 5: 49–51. https://doi.org/10.1080/21507740.2013.863252.

46. Müller, Sabine, and Henrik Walter. 2014. Neither Speculative nor Narrow-Minded Ethics is Needed for Optogenetics-Based DBS in Psychiatry and Neurology. *AJOB Neuroscience* 5: 12–14. https://doi.org/10.1080/21507740.2014.911780.

47. Gilbert, Frederic, Alexander R. Harris, and Michael Kidd. 2021. Burnt in Your Memory or Burnt Memory? Ethical Issues with Optogenetics for Memory Modification. *AJOB Neuroscience* 12: 22–24. https://doi.org/10.1080/21507740.2020.1866104.

48. Lavazza, Andrea. 2021. Can Memory Make a Difference? Reasons for Changing or Not Our Autobiographical Memory. *AJOB Neuroscience* 12: 38–40. https://doi.org/10.1080/21507740.2020.1866106.

49. Bublitz, Jan Christoph, and Dimitris Repantis. 2021. Memory, Authenticity, and Optogenethics. *AJOB Neuroscience* 12: 30–32. https://doi.org/10.1080/21507740.2020.1866102.

50. Kostick, Kristin Marie, and Gabriel Lázaro-Muñoz. 2021. Neural Safeguards against Global Impacts of Memory Modification on Identity: Ethical and Practical Considerations. *AJOB Neuroscience* 12: 45–48. https://doi.org/10.1080/21507740.2020.1866111.

51. Robins, Sarah K. 2016. Optogenetics and the mechanism of false memory. *Synthese* 193: 1561–1583. https://doi.org/10.1007/s11229-016-1045-9.

52. Robins, Sarah K. 2018. Memory and optogenetic intervention: Separating the engram from the ecphory. *Philosophy of Science* 85: 1078–1089. https://doi.org/10.1086/699692.

53. Mihailov, Emilian, Alexandra Zorila, and Cristian Iftode. 2021. Taking Relational Authenticity Seriously: Neurotechnologies, Narrative Identity, and Co-Authorship of the Self. *American Journal of Bioethics Neuroscience* 12: 35–37. https://doi.org/10.1080/21507740.2020.1866105.

54. Mackenzie, Catriona, and Mary Walker. 2015. Neurotechnologies, personal identity and the ethics of authenticity. In *Springer Handbook of Neuroethics*, ed. Jens Clausen, and Neil Levy, 373–92. Dordrecht: Springer.

55. Erler, Alexandre. 2021. Optogenetic Memory Modification and the Many Facets of Authenticity. *AJOB Neuroscience* 12: 40–42. https://doi.org/10.1080/21507740.2020.1866108.

56. Lindemann, Hilde. 2001. *Damaged Identities, Narrative Repair*. New York: Cornell University Press.

57. Baylis, Françoise. 2013. "I Am Who I Am": On the Perceived Threats to Personal Identity from Deep Brain Stimulation. *Neuroethics* 6: 513–526. https://doi.org/10.1007/s12152-011-9137-1.

58. Schechtman, Marya. 2010. Philosophical reflections on narrative and deep brain stimulation. *The Journal of Clinical Ethics* 21: 133–139.

59. Zawadzki, Przemysław. 2021. Dimensions of the Threat to the Self Posed by Deep Brain Stimulation: Personal Identity, Authenticity, and Autonomy. *Diametros* 18: 71–98. https://doi.org/10.33392/diam.1592.

60. DeGrazia, David. 2005. *Human identity and bioethics*. New York: Cambridge University Press.

61. Schechtman, Marya. 2014. *Staying alive: Personal identity, practical concerns, and the unity of a life*, New York: Oxford University Press.

62. Schechtman, Marya. 1996. *The constitution of selves*. New York: Cornell University Press.

63. Lindemann, Hilde. 2014. *Holding and letting go: The social practice of personal identities*. Oxford: Oxford University Press.

64. Baylis, Françoise. 2012. The self in situ: A relational account of personal identity. In *Being relational: Reflections on relational theory and health law,* ed. Jocelyn Downie and Jennifer J. Llewelyn, 109–131. Vancouver: UBC Press.

65. Nyholm, Sven, and Elizabeth O'Neill. 2016. Deep Brain Stimulation, Continuity over Time, and the True Self. *Cambridge Quarterly of Healthcare Ethics* 25: 647–658. https://doi.org/10.1017/S0963180116000372.

66. Erler, Alexandre, and Tony Hope. 2014. Mental Disorder and the Concept of Authenticity. *Philosophy, Psychiatry, & Psychology* 21: 219–232. https://doi.org/10.1353/ppp.2014.0032.

67. Erler, Alexandre. 2012. One Man's Authenticity is Another Man's Betrayal: A Reply to Levy: One Man's Authenticity is Another Man's Betrayal. *Journal of Applied Philosophy* 29: 257–265. https://doi.org/10.1111/j.1468-5930.2012.00562.x.

68. De Brigard, Felipe. 2014. Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese* 191: 155–185. https://doi.org/10.1007/s11229-013-0247-7.

69. Loftus, Elizabeth F., and Jacqueline E. Pickrell. 1995. The formation of false memories. *Psychiatric Annals* 25: 720–725. https://doi.org/10.3928/0048-5713-19951201-07.

70. Schechter, Elizabeth. 2018. *Self-consciousness and "split" Brains: The Minds' I*. Oxford: Oxford University Press.

71. Walker, Mary Jean. 2012. Neuroscience, self-understanding, and narrative truth. *AJOB Neuroscience* 3: 63–74. https://doi.org/10.1080/21507740.2012.712603.

72. Van Leeuwen, D.S., and Neil. 2007. The spandrels of self-deception: Prospects for a biological theory of a mental phenomenon. *Philosophical Psychology* 20: 329–348. https://doi.org/10.1080/09515080701197148.

73. Von Hippel, William, and Robert Trivers. 2011. The evolution and psychology of self-deception. *Behavioral and Brain Sciences* 34: 1–16. https://doi.org/10.1017/S0140525X10001354.

74. Lauria, Federico, Delphine Preissmann, and Fabrice Clément. 2016. Self-deception as affective coping. An empirical perspective on philosophical issues. *Consciousness and Cognition* 41: 119–134. https://doi.org/10.1016/j.concog.2016.02.001.

75. Dworkin, Gerald. 2015. The nature of autonomy. *Nordic Journal of Studies in Educational Policy* 2015 (2): 28479. https://doi.org/10.3402/nstep.v1.28479.

76. Kühler, Michael, and Nadja Jelinek. 2012. *Autonomy and the self*. Dordrecht: Springer.

77. Mackenzie, Catriona. 2008. Introduction: Practical identity and narrative agency. In *Practical identity and narrative agency*, ed. Catriona Mackenzie, and Kim Atkins, 1–28. London: Routledge.

78. Mackenzie, Catriona, and Natalie Stoljar. 2000. *Relational autonomy: Feminist perspectives on autonomy, agency, and the social self*. Oxford: Oxford University Press.

79. Oshana, Marina. 1998. Personal Autonomy and Society. *Journal of Social Philosophy* 29: 81–102. https://doi.org/10.1111/j.1467-9833.1998.tb00098.x.

80. Oshana, Marina. 2016. *Personal Autonomy in Society*. London: Routledge. https://doi.org/10.4324/9781315247076.

81. Oshana, Marina. 2020. Relational autonomy. In *International encyclopedia of ethics*, 1–13. Wiley Online Library. https://doi.org/10.1002/9781444367072.wbiee921.

82. Bravo-Rivera, Christian, and Francisco Sotres-Bayon. 2020. From isolated emotional memories to their competition during conflict. *Frontiers in Behavioral Neuroscience* 14: 36. https://doi.org/10.3389/fnbeh.2020.00036.

83. McAdams, Dan P., and Benjamin K. Jones. 2017. Making meaning in the wake of trauma. In *Reconstructing meaning after trauma*, ed. Elizabeth Altmaier, 3–16. London: Elsevier. https://doi.org/10.1016/B978-0-12-803015-8.00001-2.

84. Wilson, John P. 2007. *The posttraumatic self: Restoring meaning and wholeness to personality*. New York: Brunner-Routledge.

85. Lindemann, Hilde. 2019. *An invitation to feminist ethics*. New York: Oxford University Press.

86. Delker, Brianna C., Rowan Salton, Kate C. McLean, and Moin Syed. 2020. Who has to tell their trauma story and how hard will it be? Influence of cultural stigma and narrative redemption on the storying of sexual violence. *PLOS ONE* 15(6): e0234201. https://doi.org/10.1371/journal.pone.0234201.

87. Brunet, Alain, Joaquin Poundja, Jacques Tremblay, Eric Bui, Emilie Thomas, Scott P. Orr, Abdelmadjid Azzoug, Philippe Birmes, and Roger K. Pitman. 2011. Trauma reactivation under the influence of propranolol decreases posttraumatic stress symptoms and disorder: 3 open-label trials. *Journal of Clinical Psychopharmacology* 31: 547–550. https://doi.org/10.1097/JCP.0b013e318222f360.

88. Brunet, Alain, Daniel Saumier, Aihua Liu, David L. Streiner, Jacques Tremblay, and Roger K. Pitman. 2018. Reduction of PTSD Symptoms With Pre-Reactivation Propranolol Therapy: A Randomized Controlled Trial.

*American Journal of Psychiatry* 175: 427–433. https://doi.org/10.1176/appi.ajp.2017.17050481.

89. Brunet, Alain, Émilie. Thomas, Daniel Saumier, Andrea R. Ashbaugh, Abdelmadjid Azzoug, Roger K. Pitman, Scott P. Orr, and Jacques Tremblay. 2014. Trauma Reactivation Plus Propranolol Is Associated With Durably Low Physiological Responding During Subsequent Script-Driven Traumatic Imagery. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie* 59: 228–232.

90. Kindt, Merel, and Arnold van Emmerik. 2016. New avenues for treating emotional memory disorders: Towards a reconsolidation intervention for posttraumatic stress disorder. *Therapeutic Advances in Psychopharmacology* 6: 283–295. https://doi.org/10.1177/2045125316644541.

91. Brunet, Alain, Scott P. Orr, Jacques Tremblay, Kate Robertson, Karim Nader, and Roger K. Pitman. 2008. Effect of post-retrieval propranolol on psychophysiologic responding during subsequent script-driven traumatic imagery in post-traumatic stress disorder. *Journal of Psychiatric Research* 42: 503–506. https://doi.org/10.1016/j.jpsychires.2007.05.006.

92. LeDoux, Joseph E. 2000. Emotion Circuits in the Brain. *Annual Review of Neuroscience* 23: 155–184. https://doi.org/10.1146/annurev.neuro.23.1.155.

93. Mahr, Johannes B., and Gergely Csibra. 2018. Why do we remember? The communicative function of episodic memory. *Behavioral and Brain Sciences* 41: 1–93. https://doi.org/10.1017/S0140525X17000012.

94. Lavazza, Andrea. 2015. Erasing traumatic memories: When context and social interests can outweigh personal autonomy. *Philosophy, Ethics, and Humanities in Medicine* 10: 3. https://doi.org/10.1186/s13010-014-0021-6.

95. Hobfoll, Stevan E., Allison E. Gaffey, and Linzy M. Wagner. 2018. PTSD and the influence of context: The self as a social mirror. *Journal of Personality* 1: 76–87. https://doi.org/10.1111/jopy.12439.

96. McAdams, Dan P. 2019. Psychopathology and the self: Human actors, agents, and authors. *Journal of Personality* 1: 146–155. https://doi.org/10.1111/jopy.12496.

97. Frankfurt, Harry G. 1971. Freedom of the will and the concept of a person. *The Journal of Philosophy* 68: 5–20. https://doi.org/10.2307/2024717.

98. Frankfurt, Harry G. 1988. *The importance of what we care about: Philosophical essays*. New York: Cambridge University Press.

99. Maslen, Hannah, Jonathan Pugh, and Julian Savulescu. 2015. The Ethics of Deep Brain Stimulation for the Treatment of Anorexia Nervosa. *Neuroethics* 8: 215–230. https://doi.org/10.1007/s12152-015-9240-9.

100. Dranseika, Vilius. 2021. Authenticity, Self-Defining Memories, and the Direction of Change. *AJOB Neuroscience* 12: 48–49. https://doi.org/10.1080/21507740.2020.1866112.

101. Westlund, Andrea C. 2009. Rethinking relational autonomy. *Hypatia* 24: 26–49.

102. DeScioli, Peter, and Robert Kurzban. 2009. The alliance hypothesis for human friendship. *PLoS ONE* 4(6). https://doi.org/10.1371/journal.pone.0005802.