

NEIL LEVY

CHANGING ONE'S MIND.

THE ETHICS OF MEMORY ERASURE IN *ETERNAL SUNSHINE OF THE SPOTLESS MIND**

1. Intro 2. Memory matters 3. Harm to Self / Harm to Others 4. Conclusion

1. Intro

In *Eternal Sunshine of the Spotless Mind*, a company called Lacuna (the word means “gap” or “blank space”) offers its clients the service of erasing unwanted memories. Clementine, Joel's former girlfriend,



Monica Rabà dove sono le cose?

decides to take advantage of Lacuna's services, to have their relationship erased from her mind. When Joel realizes what she has done, he decides to emulate her, though he come to regret the decision, and resists the procedure. The scenario is intriguing; many of us are fascinated by the power of memory to shape our view of the world and of ourselves. Indeed, *Eternal Sunshine* is merely the latest in a long line of films to explore the power of memory: from Hitchcock's *Marnie* (1964), in which the title character represses the memory of a murder but remains haunted by its

* *Eternal Sunshine of the Spotless Mind*, film diretto da Michel Gondry nel 2004, interpretato da Jim Carrey e Kate Winslet, è uscito in Italia con il titolo *Se mi lasci ti cancello* [NdR].

profoundly disturbing consequences, to more recent science fiction films in which memories are implanted, such as *Total Recall* (1990) and *Dark City* (1998), as well as films exploring the world of the amnesic, such as *Memento* (2000) and *50 First Dates* (2004). These films grip us because we know that in some important sense our memories are us; to lose one's personal past is almost like dying, and to have it altered without our consent would be the gravest kind of assault. But why does memory matter so much? Why would it be wrong for others to alter our memories without our consent? Is such memory alteration technically feasible? And if the technology to change or erase memories were to be developed, would we, *should we*, use it? This paper aims to explore these questions.

2. Memory matters

Why do our memories matter so much to us? John Locke, the great English philosopher of the seventeenth century, is often interpreted as arguing that memory constitutes personal identity; so long as a person can remember past experiences, she remains the same person as the person who had those experiences. This so-called memory criterion of personal identity has come under a lot of fire from philosophers who claim that it is circular: since I can only remember what has actually happened to me, memory *presupposes*, and therefore cannot *constitute*, personal identity. However, it is clear that the claim that memory constitutes personal identity is importantly right. Critics of Locke may be right in claiming that the memory criterion cannot do the work demanded of it - it cannot provide us with a means of saying whether two individuals, at different times, are in fact one and the same person - but memory is nevertheless essential to our sense of who we are. In daily life, and in much of moral philosophy as well, we are simply not concerned with the question of personal identity in the sense in which Locke's criterion allegedly fails. Instead, we are concerned with identity in the

sense of who we are and where we stand; what matters to us, and memory is essential for *this* sense of identity.

Consider the phenomenon of what is sometimes called an *identity crisis*. Someone who experiences such a crisis is not puzzled as to whether they are, as philosophers say, numerically identical to themselves at some previous time. They do not wonder whether the person who got out of bed this morning is literally the same person as the one who went to sleep last night (“maybe I’ve been kidnapped and replaced by a replica”). Instead, they wonder whether the person they are has a firm grip on his or her values; whether they are making their own way in the world or living out someone else’s (parents, community, authority figures) idea of who they should be; they might feel estranged from the values they espouse and the life they are leading. At the heart of all this is memory. Someone who cannot recall the relationships in which they are embedded, the projects in which they are engaged, and the values they pursue cannot experience an identity crisis, because they have too little in the way of identity. The complete amnesic cannot be alienated from her identity because alienation is a sense of distance from something we can see and understand, not from something that has disappeared from our world.

When people are asked who they are, they naturally respond in memory-dependent terms. They may mention their job, their family, their nation. None of these are things they carry around with them. Instead, they *are* these things insofar as they represent patterns of interaction and activity which extend from the past into the future. We are also strongly marked by significant events in our lives, even if they leave no physical scars. The man or woman who was in the World Trade Centre on September 11 will be marked by those events for life (indeed, even those of us who only watched events unfold on television are profoundly altered by them). Public and private events unfold and leave those who participate altered by their passing, for better or for worse. We

do not live moment-by-moment; instead, we live by connecting our past to our future, and our memories are the essential thread by which we make these connections. This is why Alzheimer's is such a cruel disease: to watch someone gradually lose their memories to dementia is to watch them almost literally unravel before your eyes.

These are some of the reasons why memory matters so much to us, why we are fascinated by its loss, and by the possibility of altering it, augmenting it, or erasing it. Memory erasure of the kind imagined in *Eternal Sunshine* is unlikely to be feasible anytime soon. Indeed, some neuroscientists wonder whether it will ever be possible. Part of the problem is that memories are not stored in one place, where they can easily be accessed and altered. Instead, at least once the memories have been consolidated (moved from short-term to long-term memory) they are stored in widespread cortical networks. Worse, so far as the prospect of targeting individual memories is concerned, the neurons involved in storing one memory (say, Joel's memory of meeting Clementine) may also be involved in different memories, or perhaps different functions altogether, as well.

However, there is one line of research which might offer us the prospect of targeting individual memories for erasure. According to the *reconsolidation hypothesis*, memories are not merely recalled from storage whenever they are used, in the way that a computer recalls memory from its hard drive. Instead, memory is a recreative process; when you recall a memory, you rewrite it. Now, memories are particularly vulnerable to erasure when they are new, in part because the move from short-term to long-term storage can be disrupted. Football players who experience head trauma during a match may be able to recall the plays perfectly for the few seconds they need them in working memory, but their injury may prevent the memory being consolidated, and the player may not be able to recall the match at all the next day. Some researchers

have suggested that we can take advantage of this temporary vulnerability of memory: we can provoke the memory, requiring it to be reconsolidated, and then disrupt the process¹. Theoretically, this might result in the memory being erased. Perhaps Charlie Kauffman, the writer of *Eternal Sunshine*, is aware of this research, because the procedure he imagines is consistent with it. Joel is asked to recall in detail his memories of Clementine, perhaps in order to provoke reconsolidation².

There is no doubt that recall is at least partially a recreative process. The memories we recall are influenced by the goals we have at the moment of recollection, our intervening experiences and our reinterpretations³. However, it is very controversial whether disrupting reconsolidation could have far-reaching effects on memory, and even if it could, it seems unlikely that a memory as significant to our personal biographies as a love affair could be erased. In a well-known essay⁴, Dan Dennett asks us to imagine what would be involved in inserting a belief in the brain by using a technology which is the mirror-image of Lacuna's memory erasure machine - "brain writing". Supposed we wanted to insert into a subject the false belief "I have an older brother living in Cleveland". Dennett points out that either the technique will fail, or we shall end up with an entirely irrational subject, because beliefs do not generally come in all by themselves. Instead, they come in clusters, and the more central to our identity (in the sense in which I am using the word here) the larger the cluster. So wiring in the belief that Tom has an older brother in Cleveland requires wiring in a rather large cluster of beliefs: that he played with his brother as a small boy; that they

¹ S. J. Sara, *Retrieval and Reconsolidation: Toward a Neurobiology of Remembering*, in «Learning & Memory», 7, 2000, pp. 73-84.

² S. Johnson, *The Science of Eternal Sunshine*, in «Slate», March 22, 2004, www.slate.com/id/2097502/.

³ D. L. Schacter, *Searching for Memory: The Brain, the Mind, and the Past*, Basic Books, New York 1996.

⁴ D. Dennett, *Brain Writing and Mind Reading*, in *Brainstorms: Philosophical Essays on Mind and Psychology*, Penguin Books, London 1978.

went to the same school (or that they went to different schools); that their parents favored one or the other, or at any rate treated them differently; beliefs about what he is doing, whether he is married, how popular he was, whether he overshadowed Tom or helped him, and so on; the list can be carried on indefinitely. Central experiences and relationships are not isolates in our mental economy; instead, they spread their shadow over almost all of our lives.

Now consider Joel's life with Clementine. Erasing her from his mind would require almost as extensive a rewiring of his brain as inserting a belief about a brother. It would require not only erasing her from his mind, but also all the experiences that they had together, and which her presence colored. Clementine's presence will have had literally innumerable effects on Joel's mental life for the period they were together, some subtle, some less so. Many of the activities in which he engaged were *joint* activities: dancing, talking, having sex, but also, more subtly, going to a restaurant to eat Chinese. Even activities in which they merely engage in parallel, as it were, will have been colored by the presence of the other. Going to the movie to see *Eternal Sunshine of the Spotless Mind* is a somewhat different experience depending upon whether you go to see it with your brother, your lover or a friend. Their presence inflects the way you experience it, and afterwards you shall probably talk about it. The manner in which the experience will be moved from short to long-term memory will almost certainly be significantly modulated by the way you encode it and the manner in which you think about it, as you discuss it afterwards⁵. It will be a different experience. Joel's memories of Clementine cannot easily be extracted from his mind without unravelling many other threads. It would be relatively easy - if the technology existed - to insert an isolated and inconsequential memory in someone's brain - say, a memory of

⁵ D. L. Schacter, *op. cit.*

having found a quarter. And it would be relatively easy to erase an isolated memory. But erasing a central, identity-defining relationship may well be beyond us, not for technological reasons, but for reasons to do with the structure of experience.

3. Harm to Self / Harm to Others

Set all that aside. Suppose, for the sake of argument, that memory erasure on the scale *Eternal Sunshine* depicts will one day be possible. Were that the case, should we use the technology? An ethical assessment of the technology will have to consider two main questions: will it harm the user? Will it harm others?

How might memory erasure or alteration result in harm to self? Recall that our identity, in the sense of who we are, what matters to us and what we stand for, is tied closely to our recall of significant moments from our lives. As Marya Schechtman points out, persons are essentially narrative beings: they are constituted by the stories they tell themselves and one another about their lives⁶. This is the case both at the most trivial and the most profound level of our lives. Almost all of our daily activities take their significance from their place in an ongoing narrative. I type this sentence *in order to* finish this paragraph, and I aim to finish this paragraph *in order to* finish this paper; and I want to finish this paper *in order to*... what? We can trace the significance of this work to me ultimately all the way to what Sartre called my fundamental project, which is my way of living my life⁷. Perhaps I aim (at a level slightly less fundamental than the one Sartre had in mind) to boost my CV and get a better job, thereby to increase my income and my personal comfort, or perhaps I aim at recognition from my peers, or at increasing the amount knowledge in the world. In any case, the meaning of my action is

⁶ M. Schechtman, *The Constitution of Selves*, Cornell University Press, Ithaca 1996.

⁷ J.-P. Sartre, *Being and Nothingness* (1943), eng. tr. Washington Square Press, New York 1956.

constituted in very important part by the threads which link my past to my future.

Now, as Schectman recognizes, none of our personal narratives are wholly consistent or coherent. Nor are they ever wholly true; each of us is likely to misremember certain events. A certain amount of incoherence need not matter, nor a certain amount of falsity. It is a difficult matter to identify the point at which incoherence or falsity ceases to be innocuous. Part of the problem comes from the fact that more than one kind of good can depend upon the kinds of narratives we construct for ourselves, and these goods can conflict. Narratives are routes to self-knowledge, which is a good that is - arguably - intrinsic, that is, valuable in its own right, as well as a good that is instrumentally valuable inasmuch as it allows us to achieve other goods. But narratives can also be instrumentally valuable independently of their truth.

An example will make this clearer. Owen Flanagan, a well-known American philosopher, recounts of a small part of his own biography to illustrate how memory can be instrumentally valuable, independent of its truth⁸. He had, he tells us, very few friends as a young child. But he did have one close friend, Billy, with whom he spent many happy hours playing. Later he lost touch with Billy, but the memory of this important friendship gave him the confidence he needed to approach people and make new friends. Years later he discovered that his Billy memories were almost entirely false. Billy had been the son of one's of his father's colleagues, and had visited the Flanagans just once. Owen had indeed played with Billy, but they had never been friends. The close and long-lasting friendship was a fabrication, built upon the flimsiest of foundations - yet the influence of the 'memory' had been very real. It really had contributed significantly to Owen's later success at making friends.

⁸ O. Flanagan, *I remember you*, in *Self-Expressions: Mind, Morals, and the Meaning of Life*, Oxford University Press, New York 1996, pp. 88-98.

Owen's memory was false, and therefore could not contribute to his self-knowledge. Yet it was instrumentally valuable, helping him to achieve goods that mattered to him. More usually, we can expect instrumental value and intrinsic value to coincide, for the following reason: generally, being able to achieve the goods that matter to us depends on having true beliefs, both about ourselves and about the world. If you want to avoid danger, it helps to be able to accurately distinguish tigers from trees, and to know how good you are at outrunning the former and climbing the latter. If Owen had no social skills and no capacity to develop them, his false memories would not have been instrumentally valuable; it is only because - by chance - his false memories were not an inaccurate guide to his capacities that they proved useful.

The more central the capacities to the agent's sense of herself, and the more central to the projects she undertakes, the more important it is for her to have an accurate sense of them. Owen went on to become an important philosopher of mind; it was therefore important to him to have a good sense whether he was more talented intellectually or on the sporting field. For most of us, the kind of work we do is central to our sense of identities, in the sense with which we are here concerned. Friendship, family and relationships are also central to this sense of identity, and we shall therefore be concerned to have true beliefs with regard to these. It does not matter very much whether each of our memories regarding our interactions with those closest to us are accurate, so long as our general sense of the shape of the relationship - of its narrative course - is accurate⁹. Reminiscing with family and friends we occasionally discover events concerning which we have divergent memories. It doesn't matter much, so long as each of us is correct in their general sense of the importance we have for each other, the place we each occupy in the others' lives and their affections.

⁹ M. Schechtman, *op. cit.*

This suggests one important reason why erasing certain memories can constitute a significant harm to oneself. Some of the memories we might be tempted to erase are relatively inconsequential: all those petty embarrassments and humiliations which haunt each of us. But some of our memories, including some of our most painful, are important guides to our abilities and limitations. They constitute self-knowledge, and this self-knowledge is – at least – instrumentally valuable to us. Consider Joel and Clementine once more. When they meet again, after having their respective memories erased, each is drawn to the other. Now, there is more than one possible explanation of what is occurring here. One possibility is that some memory of each other survives Lacuna's memory erasure. Joel, after all, has resisted the procedure. But Clementine did not resist, and she seems equally attracted to Joel as he is to her. Another explanation is that Lacuna's technique targets *semantic* and *episodic* memory only. Semantic memory is the memory we have for facts: Paris is the capital of France, mozzarella is a kind of cheese, Billy is my best friend. Episodic memory is the memory we have for particular kinds of events in our life: that time I ate mozzarella with Billy in Paris, that time Billy and I went sledding. Now, in addition to semantic and episodic memory, there are a number of other memory systems, but the one that might be relevant here is *implicit* memory. Our implicit memory is our memory for facts, events, procedures, which bypasses consciousness: it is the kind of memory which we have (normally) without being aware of it.

Consider, for instance, the famous case of HM (subjects are identified by initials alone in the medical and neuroscientific literature, in order to protect their privacy). HM suffered from extremely severe epilepsy, experiencing seizures of such an intensity and frequency that his doctors believed that brain surgery was required to treat him (such operations have frequently allowed sufferers from the most severe forms of epilepsy to resume

a normal life). Unfortunately, the surgery was spectacularly unsuccessful: not only did it fail to prevent his seizures, it left HM with severe amnesia. HM had the operation in 1953; ever since, he has been stuck in the present moment. He suffers from *anterograde amnesia*, which means that he is unable to learn anything new. He does not recognize anyone he met only after the surgery and he cannot learn most new skills. His explicit memory is, for all intents and purposes, profoundly dysfunctional. But HM still has some implicit memory intact. For instance, his skills at video games, upon which he has been tested, gradually improves with practice. Day after day his performance on such games improves. Yet each time he takes himself to be playing the game for the first time! In addition, sufferers from profound anterograde amnesia may exhibit *priming*: though they cannot recall being shown a list of words, nevertheless they perform much better at word completion tasks than they would have if they had not previously been exposed to the list. Their memories for the list, like HM's memory for the game is implicit: it affects their behavior, but they are unable consciously to access it.

There is evidence that normal people have implicit - as well as explicit - memory for faces. This implicit memory is dissociable - that is, come apart - from our explicit memory, in surprising ways. There are two neurological disorders which may be caused by such dissociations. In *prosopagnosia*, subjects fail to (explicitly) recognize faces, but they still show signs of implicit memory for them. For instance, they show skin conductance responses to familiar faces (skin conductance is a measure of how well our skin conducts electricity; it is the physiological reaction that is measured by lie detector tests). In *Capgras's delusion*, subjects explicitly recognize faces, but they do not show any skin conductance response to them. As a result of the loss of a sense of familiarity, sufferers from the delusion think that the people they recognize are merely replicas of their

friends and family; perhaps aliens or robots who have taken on their identity. They conclude that the body that *looks* like their wife or their father can't *be* the familiar person, because they do not get that feeling of warmth upon seeing them¹⁰.

So one possibility is that Joel and Clementine have an implicit memory of each other. In that case, they would have a sense of familiarity, though they would not recognize one another; their experience would be analogous to that of sufferers from prosopagnosia. Another possibility is that, even though they are very different from one another, each is somehow the other's 'type'. They are the kind of person they find attractive, or, at least, have elements of personality or physical style which draws the other. Whichever hypothesis is true, their memory erasure puts them at an important disadvantage. They cannot learn from their failed relationship. To be sure, in the course of the film each learns that they had a relationship, and hears a record of the kinds of things they disliked about one another. But this second-hand information is unlikely to prove useful to them in the way that a direct memory of the course of their relationship might. If they are to avoid putting each other through the same pain once again, they need to learn from what went wrong. It is possible that recalling their mistakes will enable them to avoid repeating them. It is also possible that they cannot avoid their mistakes; that though the relationship is worth having they cannot hope to make it last. Nevertheless, this too is knowledge worth having. At the end of the movie, Joel and Clementine step back onto the merry-go-round. We get the sense that their relationship is doomed once more. Yet perhaps they make the right decision: having the relationship, appreciating the moments of joy it allows them, and learning about themselves and about others through it, might make the pain it causes them worthwhile. But it can only be worthwhile

¹⁰ H.D. Ellis, A.W. Young, *Accounting for delusional misidentifications*, in «British Journal of Psychiatry», 157, 1990, pp. 239-248.

if the self-knowledge and the appreciation of the goods the relationship can provide them with can survive its end. If they cannot recall, they are at the mercy of the forces that, through implicit memory or merely attraction, impel them toward one another. As the American philosopher George Santayana once remarked, «Those who cannot remember the past are condemned to repeat it».

One important reason why we ought to refrain from erasing our memories, were we to acquire the power, is that doing so would impoverish our self-knowledge, and thereby put at risk our capacity to pursue the goods we value. However, it may be that memory erasure will harm others, and not only the person whose memory is at stake. Showing that this might be so is important, because considerations of autonomy normally trump considerations of self-harm. That is, if my actions will harm no one but myself, then it would normally be wrong for others to coerce me into refraining from them. I have the right to harm myself. But I do not have the right to harm others without their consent, If, therefore, memory erasure might harm others, we might be permitted to prevent its use.

How might memory erasure harm others? In several ways. Our narratives, which form the core of our identities in the sense with which we are here concerned, are not merely personal and private stories. Each of us is perpetually at risk of what psychologists call *confabulation*: inventing more or less plausible stories that bear little relationship to reality to explain what we do and why. Without a public check on what happens to us, the risk that we shall slip into unchecked fantasy is high. Once again, the way memory works is important to this process. *Source memory*, our memory for where we acquired a piece of information, is dissociable from semantic memory¹¹. This fact has led to a number of real-life confusions and mistakes: for instance,

¹¹ D. L. Schacter, *op. cit.*

misidentifications of crime suspects by eyewitnesses, who correctly recall seeing the face before them (perhaps in a book of mug shots) but do not recall the source of their memory and therefore infer that the person is the perpetrator. The dissociation of source memory from semantic memory explains how people can perfectly innocently come to believe their own fantasies, and this, too, has had tragic real-world consequences. It is therefore very important, for our self-knowledge, that we associate with people who can corroborate important elements of our life story.

Moreover, even if we succeed in retaining accurate memories of the events in our lives, we would likely be damaged by our inability to share them with those who feature centrally in its unfolding narrative. In the past decade, philosophers have devoted a great deal of effort to understanding the concept of *recognition*, first introduced into philosophy by the great German philosopher Hegel. Hegel saw that it was very important for us social animals that our worth is recognized by other people. But, as he also saw, our sense of self-worth depends upon recognition from others that we ourselves think worthy of recognition in turn; moreover, the recognition must be freely given if it is to be valuable. The absence of this kind of recognition is, as more recent philosophers have emphasized, profoundly damaging to our identity (once again, in the sense of identity with which we are here concerned).

Why should recognition be important to our identity? As Charles Taylor points out¹², this is a consequence of the extent to which our identity is *dialogical*: we understand ourselves in terms which we fashion in dialogue with others. Whereas other animals have (at best) only a very rudimentary culture, which does not significantly shape them, we humans are essentially cultural

¹² C. Taylor, *The Politics of Recognition*, in *Philosophical Arguments*, Harvard University Press, Cambridge, Mass. 1994, pp. 225-256.

animals, and culture is by definition something which exists only intersubjectively: as a result of the interaction of human beings. Our identity is profoundly cultural, and it is therefore up for negotiation and renegotiation in the stories we tell one another. Of course, as we mature we internalize this story-telling; we begin to construct our narratives for ourselves. But the extent to which, even today in cultures which place a high value on the autonomy of the individual, we can ever entirely break out of the dialogical mode is limited. We engage in lifelong conversations with others – actual conversations, with those from whom we seek recognition, and internal conversation, with those who have passed from our lives.

Now, as profoundly damaging as a lack of recognition might be, how much worse is the failure, not to recognize one's *worth*, but even to acknowledge that one has played a significant role in one's life? We cannot demand recognition of others: if relationships fail, we cannot prevent our lovers from walking out. We do not have a right to their time and affection, nor even to their attention. Sometimes, we must move on. But we can reasonably expect that our former lovers will at least acknowledge that we once *had* a relationship. We sometimes advise friends to try to forget their mistakes, to put them entirely behind them. In the worst cases, we might mean it literally: one might do better to entirely forget the abusive relationship (so long as they memory will not help us to avoid a repetition). But in more run-of-the-mill cases, the advice is more metaphorical. Though we know it is hard to pull it off, we often think that in acknowledgement of the significance of the relationship ex-lovers ought to remain friends, or at least cordial to one another. Of course, often this is not psychologically possible – when the relationship is broken off acrimoniously, perhaps after bad behaviour by one or the other. But even in these cases, the coldness with which the former lovers might treat one another when they meet is a form of

recognition. It has been said that hate is not the opposite of love, since both are forms responding to the individuality of the other. The profound indifference that is the product of true forgetting, however, might be far more damaging to our sense of worth than mere hate, or the lack of recognition. Just consider how Joel responds when Clementine cannot recall even the existence of their relationship.

Why is this profound indifference so damaging to our sense of identity? When we have had a long and significant relationship with another person, our identity-constituting narrative now incorporates that relationship into its core. The relationship might be conceptualized as the culmination of an important plot-line, as it were, in the way in which relationships are often understood in many movies. If the other person then erases the memory of that relationship from her mind, there is an important sense in which it is as if it never was. There is little practical difference, for us, between actually having had the relationship and merely having fantasized it (in this regard, it is very important whether or not third-parties recall the relationship, as *Eternal Sunshine* recognizes. If not only Clementine, but all their mutual friends, have forgotten the relationship, then it truly cannot serve any identity-constituting role any longer. If, on the other hand, Clementine has erased it but everyone else remembers it, then the effect is closer to the kind of result caused by the death of the partner: the dialogue can continue, though it can no longer continue with the most central actor in it).

Now, it is one thing to say that erasing a memory would cause a harm to other people, and quite another thing to say that it is – all things considered – impermissible to erase the memory. There are many things that are morally wrong, but which we are permitted to do. We may, for instance, lie to each other, without fear of legal sanction. Roughly speaking, the limits of what we are

permitted to are defined by John Stuart Mill's *harm principle*¹³. The harm principle simply states that each of us has the right to act as we like, so long as our conduct does not result in harm to (non-consenting) others. According to Mill, no one can legitimately prevent others from doing anything which does not infringe the harm principle; we cannot prevent others from doing things just because we find them distasteful, or even because we believe (or *know*) that they will later regret doing them. We may advise and persuade, but we cannot coerce, except when harm is threatened.

Applying Mill's harm principle yields the result that memory erasure is permissible when it harms no one, or when it harms only the person who has consented to it. But, as we have just seen, memory erasure would sometimes hurt other people as well. Does that imply that, in those cases at least, it is impermissible?

I think not, not, at least, in cases like the one depicted in *Eternal Sunshine*. The harm principle captures our sense of what is permissible and what is forbidden only roughly. Think of the case of lying once more: lying to those close to you may well harm them, yet we cannot coerce adults into telling the truth (except in a judicial context). If we examine the rationale for the harm principle, we can come up with a better test for distinguishing permissible and impermissible acts.

The harm principle is designed to protect the *autonomy* of each individual. The idea is this: each of us has the right to pursue his or her own conception of the good life, without coercive interference from others. Historically, this idea emerged for merely pragmatic reasons, in response to the wars of religion that racked Europe in the wake of the Reformation¹⁴. The alternative to finding a *modus vivendi* – a means of getting along with one another – was endless and ruinous war. But by the eighteenth

¹³ J. S. Mill, *On Liberty* (1859), Penguin Books, London 1985.

¹⁴ J. Rawls, *Political Liberalism*, Columbia University Press, New York 1993.

century, the doctrine of tolerance for other ways of life was increasingly recognized as *moral* principle. We have a *right* to pursue our own conception of the good life. Part of the justification for this idea comes from political philosophers pondering the purpose of the state. Many philosophers argued that the state existed only to allow autonomous individuals to pursue their own projects; since the state is constituted by the free adhesion of individuals, its legitimacy depends upon allowing each to pursue their projects without interference. It is precisely this doctrine that is expressed in the American Declaration of Independence: each of us has the 'inalienable right' to the 'pursuit of happiness' (as each of us sees it); the end of government is to secure these rights, and that when a government 'becomes destructive of these ends, it is the right of the people to alter or abolish it'.

Autonomy, the freedom to pursue our own conception of the good life without interference, is such an important good, for us in the societies founded on these principles of liberal political thought, that a certain amount of harm to others is tolerated in its pursuit. If, in pursuing goals fundamental to my conception of the good life, I cause you harms that are relatively superficial - which do not affect your fundamental interests, do not prevent you from pursuing your conception of the good life and do not put you at an unfair disadvantage - then I do not wrong you. Some people, for instance, are offended by the sight of gay couples. But since pursuing a relationship is a fundamental part of living a good life, on most people's conception, we need a very strong reason to prevent anyone from having such a relationship. The fact that some people are offended, even physically revolted, is not sufficient reason. In general, we are allowed to cause relatively minor harms, if we have no other practical way of pursuing goals and projects central to our conception of the good life.

4. Conclusion

Given that this is the case, we cannot simply infer from the fact that memory alteration or erasure, of the kind we see in *Eternal Sunshine*, might harm other people that it should not be permitted. Though the harm that Clementine causes Joel is relatively great, it seems that it is not great enough to make the action impermissible. Joel has had an important part of his life swept away by Clementine's actions, but he remains able to pursue the projects that matter to him. In time, he is able to undo the harm (even if he does not resort to memory alteration himself). Of course, it would be hard for Clementine to make a case that erasing Joel from her memory is necessary for her to pursue her central projects (though perhaps not impossibly hard). But the fact that she acts upon her own mind raises the threshold for allowing intervention in her actions. If we have the right to a sphere of liberty, within which we are entitled to do as we choose, our brains must be included within that sphere. No one may prevent me from thinking what I like, desiring what I like or fantasising what I like; my thoughts are my affair, and no one else's. It is only when we choose to act upon them that they become, potentially, matters for the state to pronounce upon and for morality to control. Thus, I must be able to remember what I like as well. It is my mind, and my mind is me (more or less; it is at any rate an essential part of me). If my freedom is to be worth anything, it must include the liberty to change my mind, figuratively or literally. Indeed, Mill himself said as much: «Over himself, over his own body and mind, the individual is sovereign»¹⁵.

We have seen that it will often harm the person who engages in it, and sometimes may harm others as well, but that the degree and kind of harm will rarely be sufficient to justify intervening to

¹⁵ J. S. Mill, *op. cit.*, p. 69.

prevent people using the technology. Does it follow, from the fact that memory erasure will usually be permissible, that moral assessment of it is pointless? 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. We can educate and inculcate the best values. Because we decide to permit something it doesn't follow that we must approve of it, nor that disapproval may not be very effective. Ethics is concerned not only with what is permitted and what is forbidden, but also with ideals and virtues; questions of what kinds of person we should strive to be.

Moreover, given that the technology of memory erasure does not exist, morally assessing it can help guide our actions, as individuals and as citizens. Though states may not be permitted to ban memory erasure, they are not therefore required to encourage its development. States have a great deal of control over the direction of future research: they can encourage certain technologies - by direct grants to researchers, by tax breaks to industry, and so on - and discourage others, and they can justifiably do so on grounds that are wider than those mandated by the harm principle. The harm principle tells us what we *cannot* do; it doesn't tell us how to choose among the options which it does not rule out. In order to make such choices, we can appeal to wider moral principles than those of political liberalism. If we were to find that a life which used memory alteration or erasure techniques was likely to be impoverished, then we could choose not to fund its development, even if we would not be permitted to forbid its use were it developed without our help. So ethical assessment of its worth remains relevant. And of course individuals may look to such assessments in deciding whether to make use of such technologies. It is one thing whether something

should be banned; it is quite another question whether we should use it. This is one of the real values of movies like *Eternal Sunshine*: as well as entertaining, as well as exploring age-old questions – the nature of love, the place of suffering in a full life, and so on – they also allow us to explore in imagination and thereby to choose between alternative futures.

NEIL LEVY è Direttore di Ricerca dell'Oxford Centre for Neuroethics, e docente di Neuroethics al Florey Neuroscience Institutes, University of Melbourne.