

# Notre Dame Philosophical Reviews

2016.09.15

Author



**Kourken Michaelian**

## Mental Time Travel: Episodic Memory and Our Knowledge of the Personal Past

Published: September 18, 2016

**Kourken Michaelian, *Mental Time Travel: Episodic Memory and Our Knowledge of the Personal Past*, MIT Press, 2016, 291pp., \$43.00 (hbk), ISBN 9780262034098.**

**Reviewed by Robert Hopkins, New York University**

In this stimulating, widely researched but sometimes rather impressionistic book, Kourken Michaelian addresses three questions about episodic memory: what is it?, how is it able to offer us knowledge?, and why did it evolve?

Michaelian's answer to the first question is that episodic memory is one exercise of a more general capacity to represent actual and possible episodes. That capacity is also exercised in, for instance, anticipating some future happening in one's life, simulating the mental states of another, or imagining merely possible vicissitudes. As used to represent episodes from the subject's own past, this capacity turns out to be radically constructive: its representations draw on information from many sources other than the subject's past experience, to create anew a picture of how things were. The answer to the first question thus makes pressing the second: how can an operation that amounts to reconstructing or imagining the past offer us knowledge? Michaelian handles this issue by dividing it in three, but the thrust of his reply is that information from outside the subject's experience can be reliable, and subjects can sort memories from mere imaginings with sufficient reliability for episodic memory to be a trustworthy guide to her past. One element in this reply is then redeployed to tackle the third question. Part of what defines episodic memory is its distinctive subjective character. That character plays a special role in explaining how we can extract reliable information from the representations memory generates. Assuming it is advantageous to have reliable information about our personal pasts, we can thus explain why such a capacity, with its distinctive subjective character, evolved.

The book's central ambition is to offer an account of episodic memory that is consonant with contemporary psychological work on the subject. Psychology is to be our main guide in settling the nature of memory (ch.2). And, while psychology alone cannot address memory's epistemic status, Michaelian favours a naturalised approach that first concentrates on the epistemic value most amenable to empirical treatment -- reliability, and then cedes as much ground as possible to empirical work in explaining how that value is one memory realises (ch.3). Part I lays out this methodology in admirably explicit, if sometimes laboured, terms. Parts II and III form the book's core, addressing memory's nature and reliability.

In Part II, Michaelian develops his preferred account of the nature of memory, the simulation theory, by rejecting more commonsensical positions as incompatible with psychological results. The most intuitive account holds that S has an (occurrent) episodic memory of some past event  $e$  only if:

- [1] S then experienced  $e$ ,
- [2] S now has a representation of  $e$ ,
- [3] the experience in [1] and the representation in [2] are appropriately connected,

and

- [4] the content of S's current representation is the same as, or a subset of, that of her experience of  $e$ .

(This somewhat simplifies Michaelian's own exposition of the view, cf. 70.)

We might elaborate the commonsense view in various ways. In particular, we might say more about the connection in [3], requiring that it go via a continually existing 'memory trace' of the kind advocated by Martin and Deutscher in their influential paper ('Remembering' *Philosophical Review* 75: 161-96) (§5.3). And we might weaken the constraint in [4], allowing the memory representation to make explicit elements that are only implicit in S's original experience (§5.5). As Michaelian sees matters, however, neither refinement is sufficient to meet the challenge to such views from psychology. That challenge has two main elements.

the basic lesson of constructive memory research is that remembering *does* in fact routinely involve modifications that introduce information not even implicitly contained in the earlier representation. (82)

That research shows that the process of laying down and retrieving an episodic memory involves several stages: *encoding* the original experience in a short-term memory representation; *consolidating* that short-term representation into a stable long-term memory representation; the *storage* phase, in which that representation is indeed stably preserved; *retrieval*, in which it is drawn upon to generate a conscious representation of how things were; and *reconsolidation*, in which retrieval itself destabilises the long-term representation, potentially altering it so as to affect future recollection. Many of these stages influence the information available in the later, retrieved conscious representation. Some of that influence is mere selection and abstraction, consistent with [4] or its revision to allow for newly explicit information. Some, however, involves the addition of content not present in the original experience (or the representations S formed at the time). One prominent example is boundary extension, in which S's memory images represent elements of *e* that lay beyond the limits of her visual field when experiencing *e*. Another is testimonial incorporation, in which memory representations incorporate elements that S did not experience but subsequently heard about from others. A third is the phenomenon in which subjects presented with assorted words, 'recognise' some that, while not on a list presented earlier, are thematically related to words that were. A fourth is observer memories, in which S's memory of *e* presents it from a perspective other than that in the original experience, representing her 'from the outside', as just another occupant of the remembered scene.

Of course, we might refuse to classify these as cases in which the added elements are remembered, precisely because they deviate from S's original experience of the remembered event. Michaelian's interest is in developing a different line. Psychology does not view such cases as involving malfunction on the part of the memory system. Rather, they are sufficiently prevalent to be treated as normal. What if we take this idea seriously, in devising a philosophical theory of memory? How does the phenomenon appear if we attempt to accommodate the thought that these count as rememberings of the added elements, too?

Perhaps we could do this merely by weakening still further condition [4] that ties memory to earlier content. Further tinkering is, however, rendered futile by the second challenge, which forces more radical revision. There is both functional and neuropsychological evidence that episodic memory is just one exercise of a system with far wider application. Brain damage that impairs episodic memory also impairs the ability to anticipate future episodes. And the neural systems deployed in episodic remembering align closely with those used to imagine merely possible episodes, to imagine ways actual past episodes might have been, to simulate the mental states of known or unspecified individuals in certain conditions, and the like (§6.1). Michaelian's goal is to treat episodic memory as something like a natural kind, the nature of which is to be settled by empirical investigation, and in particular to individuate memory phenomena by the systems involved in their generation. Empirical work suggests we should treat episodic memory as just one application of a wider 'episodic construction system' (e.c.s.). In general, the operation of the e.c.s. does not require its output to be tied to earlier experience by either content (clause [4]) or causation (clause [3]). So why treat its use in episodic memory differently? But if earlier experience need not figure as constraining either the content or origins of memory representations, it need not figure in a definition of memory at all. (So clause [1] can go too.) Rather than attempting further revisions to the commonsense view, Michaelian instead proposes replacing it with the simulation theory (107):

S episodically remembers *e* just in case

[2] S now has a representation of *e*,

and

[5] R is produced by a properly functioning episodic construction system which aims to produce a representation of an episode belonging to S's personal past.

So much for the first question. How does episodic memory, thus understood, serve as a source of knowledge of the past? Here Michaelian is careful to limit his ambitions. He does not seek to defend episodic memory against the hyperbolic doubt of the radical sceptic. Indeed, he simply assumes that episodic memory is in fact reliable. Given that assumption, his task in Part III is only to explain how it can be, to specify the factors on which its reliability turns.

If episodic memory is to act for a subject as a reliable source of knowledge of the past, three conditions must be met. First, the information contained in the representations produced by the episodic construction system, at least when used for memory, must be reliable. Second, the subject must be able to distinguish the sources of different kinds of episodic memory -- differentiating, for example, memories of experienced events from memories of mere imaginings of possible events, and from memories of

Mental Time Travel: Episodic Memory and Our Knowledge of t... <http://ndpr.nd.edu/news/69681-mental-time-travel-episodic-mem...>  
anticipations of future events (the *source problem*). Third, she must be able to distinguish episodic memories (whatever their source) from other current products of the episodic construction system: imaginings, anticipations of future events, representations of counterfactual ways past events might have been, and so forth (the *process problem*). Michaelian takes each problem in turn.

Memory can be constructive, including incorporating content that eluded past experience, provided the new information itself stems from reliable sources. Chapter 7 argues that this will indeed be the case, for at least one such source, the testimony of others. It cites psychological evidence of 'honesty bias', our tendency both to believe what we are told and to tell what we believe to be true. This encourages optimism that the unwitting absorption of the claims of others into our episodic memories is compatible with those memories providing reliable information about remembered events. The hope is that something similar will hold for the other sources of information on which constructive memory draws.

The discussion of the source problem (ch.8) is framed in both theoretical and more concrete terms. Viewed theoretically, the issue is to construct a system that can track the source of episodic memories, so that only those originating in experiences of the world end up generating beliefs about that world. Memories don't bear explicit markers of their sources, but various factors can be combined into an effective guide to those sources. For instance, memories stemming from perceptions tend to be more detailed than those originating from waking imaginings, and memories of dreams, while also detailed, tend to fail to cohere with S's other beliefs about the world in the way memories of experiences do. Michaelian proposes that our memory system solves the source problem through a higher level element that monitors the output of the lower level episodic construction system, exploiting factors such as those just listed in order to determine whether to form beliefs about the world. That second level operation might be type 1 (quick and automatic) or type 2 (slow and reflective). Often it takes the former, sometimes the latter form. If reliability were the only desideratum, type 1 methods might be sufficient. But it is not enough for a memory system to be reliable (generating a high proportion of true beliefs), it should also exhibit power (generating true beliefs on a wide range of questions that come up -- i.e. not too often failing to generate an answer at all), and speed. The advantage of type 2 mechanisms is that they are better able to negotiate the trade-offs between these various desiderata, sacrificing some for others as context makes appropriate. Having described this solution in somewhat theoretical terms, Michaelian then sketches how a particular psychological theory of memory's operation, Mitchell and Johnson's source monitoring framework, might be seen as implementing it.

Broadly parallel moves are offered in response to the process problem (ch.9). There are no infallible guides to which process (remembering, imagining, anticipating, . . . ) the episodic construction system is at any one time implementing. Even the subject's intentions, which will often allow her to disentangle e.g. memory from imagining, are not always available, since there are unprompted memories and unintended imaginings. Nonetheless, as with sources, so with processes, there are markers that can guide our 'metacognitive' monitoring. These can be 'formal', such as the degree of control S has over the contents, once formed; content-based, such as how far the generated representation coheres with S's other representations of her world; or phenomenal, such as her sense of the position in subjective time of the represented event. Again, the monitoring might be done by systems of type 1 or 2, but it is often the former that is at work. When it is, the upshot, Michaelian suggests, is a distinctive set of 'metacognitive feelings', that suffice, not merely to distinguish episodic memories from the rest of the e.c.s.'s output, but also to distinguish each of those non-memory outputs from the others. All this is offered in the spirit of conjecture, one that future experimentation may vindicate -- though he admits that the obstacles to empirical research here are in various ways more severe than those confronting the parallel investigation of source monitoring.

The discussion of the process problem brings to the fore the theme of the subjective character of episodic memory. This theme remains prominent in the last, and shortest, of the book's four parts. Here Michaelian considers why episodic memory evolved. But now he does not individuate the phenomenon solely in terms of psychological systems. Following Tulving, he claims that human episodic memory is defined in part by its distinctive phenomenology. In particular, it centrally involves both 'chronesis', the sense of the remembered event as located in subjective time (the pastness of the remembered and the futurity of the anticipated), and 'autoindexing', the sense that the event represented intimately involves oneself. Taking these features to be partly definitive of episodic memory makes the evolutionary question harder, since many otherwise plausible accounts of why it evolved in fact target only rather weaker phenomena, such as the ability to represent 'what, when and where', or target the capacity to represent richer contents, but without any particular role for phenomenology. Michaelian sets his sights on the harder problem. There is an episodic constructive system at all because it is essential to detailed planning. Its deployment in episodic memory is an exaptation, offering subjects the further benefit of knowledge of their own past lives. And that exaptation essentially involves the chronesis and autoindexing because those phenomenal features play a key role in the solution to the process problem, and thus in enabling subjects to exploit the reliable information episodic memory proffers.

As I hope all this conveys, there is much interesting material here. Michaelian pursues his central ambition with vigour and tenacity. While he does not always offer much argument for his central claims, he does offer a rich exploration of what an account of memory would look like, when guided as far as possible by the broad contours of recent psychological work. I close by raising some

(1) The everyday notion of memory, its episodic form included, combines, as do many everyday notions, both normative and descriptive elements. Psychological research into the workings of memory, in contrast, is resolutely focused on the descriptive. This focus is inherited by Michaelian's simulation account. This is most obvious in the account's implying that one can remember an event that never occurred. Nothing in conditions [2] and [5] excludes this, and Michaelian accepts that remembering what never occurred (albeit in *misremembering*) is possible, according to his view (118). Indeed, he has earlier (§4.6) rejected any requirement that one can only remember events that really happened ('factivity'), precisely on the ground that, from the psychological point of view, the processes are the same whether initiated by the event later represented, or by some congeries of other factors. He agrees that factivity is part of the everyday conception of memory, but thinks it is something a psychologically informed theory of the phenomenon must reject. The question, though, is what relation that leaves his theory standing in to the everyday notion, and the prior philosophical attempts to analyse it. Has he simply changed topic? He notes (70) that normative constraints could simply be bolted onto the simulation theory, as extra conditions on episodic remembering. What he does not do is to consider how the two elements in such a theory, normative and descriptive, might integrate and relate; or how such integration might be essential to our everyday notion of memory and philosophical attempts to elucidate it.

(2) In fact, it is an exaggeration to say that Michaelian's own account of memory's nature is entirely cleansed of the normative. His condition [5] is supposed to distinguish episodic memory from other products of the e.c.s., including counterfactual episodic imagining about past events. But [5] can only do this if it tacitly appeals to the notion of accuracy. When one imagines a past episode as other than it actually was, one meets both [2] and [5], as stated. The difference from episodic memory is that the latter aims to represent the event as it was, whereas counterfactual imagining seeks to imagine it as it might have been (Michaelian seems close to conceding as much, 108-99). These different normative goals perhaps impose different constraints on what counts as 'proper functioning' in these two cases (a matter on which, despite occasional promises, the book is largely silent). But those differences are consequences of the more fundamental one that can only be framed in overtly normative terms. Thus Michaelian himself must distinguish various operations of the e.c.s. in terms that go beyond any psychology is interested in recognising. Again, the question arises how the normative and descriptive elements are to be made to cohere.

(3) Whether or not the process of identifying naturalistically acceptable mental kinds needs supplementing by normative considerations, the book's program dictates that descriptive psychological work take most of the strain. The refinements might include normative demands; the basics should not. However, it is not clear that Michaelian himself respects the psychology in developing the central claims of the simulation account. One of the more prominent lacunae in the book is that much of the evidence behind the two challenges from psychology to traditional philosophical accounts of episodic memory is left off the page. The evidence for both the constructive nature of episodic remembering, and for its fundamental kinship with other purported exercises of the e.c.s., such as anticipating the future, is presented relatively briefly, given its importance to the argument. (The relatively slight treatment of the latter also allows Michaelian to evade some difficult questions about how much kinship is enough to justify postulating a single system.) But, worse, not all of what is presented clearly fits Michaelian's own methodological strictures. In arguing for the constructive character of episodic memory (§5.6.1) he sometimes appeals to recognitional phenomena, such as subjects' sense that they have seen a word on an earlier occasion, or their ability to recognise people in caricatures with great ease. Recognition is no doubt a phenomenon in the same very broad area as episodic memory. But by what right does Michaelian take results involving recognition to show how episodic memory operates? Not only is recognition intuitively distinct from episodic remembering, but, more importantly, it is not even clear where to fit it into the scheme of memory phenomena he takes psychology to dictate (§2.7). Given his methodology, which requires memory phenomena to be individuated by the systems psychology discovers, some of the evidence for reconstruction does not clearly bear on the reconstructive character of episodic memory, as he identifies it.

These and other questions aside, there is a great deal here to engage and interest anyone concerned with episodic memory. Michaelian's book is a very valuable contribution to the small, but growing, philosophical literature on what is perhaps the most intriguing of the various phenomena involved in remembering.