

## **Continuities and discontinuities between imagination and memory: The view from philosophy**

Kourken Michaelian, Université Grenoble Alpes

Denis Perrin, Université Grenoble Alpes

André Sant'Anna, Massey University

**Abstract:** Though imagination and memory have much in common, philosophers of memory have so far had little to say about imagination. This has recently begun to change, as research on episodic memory as a form of imaginative mental time travel analogous to episodic future thought has threatened to undermine the view—standard in the philosophy of memory—that memory is sharply distinct from imagination. Covering a cluster of interrelated issues (including the objects of mental time travel, the reference of episodic thought, the epistemic openness of the future, the directness of our knowledge of the past, and immunity to error through misidentification in episodic memory and episodic future thought), this chapter surveys the debate between discontinuists, who argue that episodic remembering and episodic future thinking are processes of fundamentally different kinds, and continuists, who argue that the fact that they have distinct temporal orientations constitutes the only important difference between them—and hence that episodic memory is ultimately just a kind of episodic imagination.

**Keywords:** episodic imagination, episodic memory, mental time travel, continuism, discontinuism

**Number of words** - Total: 9230 (Main text: 8084 + References: 1146)

## 1 Episodic imagination and episodic memory: The continuism-discontinuism debate

This chapter surveys the debate between continuists and discontinuists about the relationship between episodic memory and episodic imagination. *Episodic memory* can most easily be characterized by contrasting it with semantic memory: whereas semantic memory is the capacity at work when one remembers that such-and-such is the case (e.g., that Ottawa is the capital of Canada), episodic memory involves mentally reconstructing an experienced episode (e.g., visiting Ottawa last year). *Episodic imagination* can similarly be characterized by contrasting it with semantic imagination:<sup>1</sup> whereas semantic imagination is the capacity at work when one imagines that such-and-such is the case (e.g., that Toronto is the capital of Canada), episodic imagination involves mentally simulating a possible episode (e.g., visiting Toronto next year). Inasmuch as they both centrally involve mental imagery—quasi-sensory experience in the absence of corresponding sensory stimulation (Nanay, 2015)—episodic memory and episodic imagination would seem, on the face of it, to be intimately related. Beyond the involvement of mental imagery, however, it is not obvious just how much overlap there is between them, and philosophers of memory, in particular, have traditionally assumed that there is in fact very little (see Bernecker & Michaelian, 2017); indeed, they have traditionally taken as one of their primary goals the demarcation of memory from imagination.

This has, however, begun to change, under the influence of impressive psychological research demonstrating that the existence of a deep difference between episodic memory and episodic imagination cannot simply be taken for granted. When Tulving (1972) first introduced the term, he defined episodic memory<sup>2</sup> as a specialized store devoted to

---

<sup>1</sup> The terminology in this area is not settled, and episodic and semantic imagination are sometimes described, for example, as sensory and suppositional imagination.

<sup>2</sup> Philosophers had traditionally referred to episodic memory as “recollective”, “experiential”, or “personal” (Brewer, 1996) memory but have, as we do here, increasingly adopted the psychological terminology.

information about the “what”, the “when”, and the “where” of experienced past events. As empirical findings indicating the existence of a tight relationship between remembering past events and imagining future events accumulated, however, Tulving (1985), along with most of his colleagues in psychology, came to adopt a definition of episodic memory as a form of *mental time travel* (MTT) in which the subject “re-experiences” past events and thus to see it as being cut from the same cloth as episodic future thought, in which the subject “pre-experiences” future events (see Suddendorf & Corballis, 2007; Michaelian et al., 2016). By emphasizing the relationship between our ability to remember the past and our ability to imagine the future, research on MTT has suggested that the difference between episodic memory and episodic imagination may not be as deep as philosophers of memory have traditionally taken it to be.

Both “continuists” and “discontinuists” (our terminology) acknowledge both similarities and differences between episodic memory and episodic future thought, but continuists maintain that, while these two forms of MTT differ *in degree*, they do not differ *in kind*; discontinuists, meanwhile, maintain that, while they may be similar in degree, they do differ in kind. Thus a discontinuist, on the one hand, might grant that the processes of remembering the past and imagining the future are executed by the same cognitive system but argue that, because remembering necessarily involves a causal connection with the relevant event, whereas imagining does not, there is nevertheless a qualitative difference between them. Discontinuism, in fact, aligns naturally with the *causal theory of memory*, according to which the difference between remembering an event and imagining it consists precisely in the presence or absence of an “appropriate” causal connection between the subject’s present representation of the event and his earlier experience of it. Martin and Deutscher’s (1966) *classical* version of the causal theory of memory understands appropriate causation in terms of the transmission of content from experience to representation via a

memory trace. More recent versions, such as Perrin's (2018) *procedural* causal theory, attempt to understand it without reference to contentful traces. The classical and procedural causal theories are discussed in section 2.

A continuist, on the other hand, might grant that imagining the future is less reliable than remembering the past but argue that this amounts to a merely quantitative difference within a single cognitive process, carried out by a single cognitive system. If discontinuism aligns with causal theories, continuism aligns with noncausal theories, such as Michaelian's (2016c) simulation theory. According to the simulation theory, episodic memory no more presupposes the presence of a causal connection between the subject's representation of an event and his experience of it than does episodic future thought and therefore ultimately reduces to one kind of episodic imagination among others: remembering is a matter of reliably imagining the past. The simulation theory is discussed further in section 2.

Perrin and Michaelian (2017) have argued that the available empirical evidence does not by itself suffice to decide between continuism and discontinuism, and our focus here will accordingly be on philosophical arguments for these positions.<sup>3</sup> We write as philosophers of memory, but both philosophers of memory and philosophers of imagination (who have been somewhat more receptive to the idea of an intimate relationship between episodic memory and episodic imagination; see Macpherson & Dorsch, 2018) have a stake in the continuism-discontinuism debate, for the appropriate relationship between their respective fields depends on how it is eventually resolved. If discontinuists are right, then philosophers of memory and philosophers of imagination may have little to learn from each other. If continuists are right, in contrast, then they likely have much to learn from each other. Accounts of the relationship between episodic memory and semantic memory (Werning & Cheng, 2017), for example, might be enriched by drawing on accounts of the relationship between episodic and semantic

---

<sup>3</sup> For additional discussion, see Gérardin-Laverge (2017) and Sant'Anna (2018).

imagination (Arcangeli, forthcoming), while recent treatments of imagination as a source of knowledge (Kind & Kung, 2016), for example, might be clarified by looking to older treatments of the epistemology of memory (Frise, 2015).

This chapter is structured around a distinction between metaphysical and epistemological varieties of continuism and discontinuism, where *metaphysical* (dis)continuism rejects (accepts) the existence of fundamental differences between episodic memory and episodic future thought, understood as cognitive processes or mental states, and *epistemological* (dis)continuism rejects (accepts) the existence of fundamental differences between the knowledge of past events and the knowledge of future events that is provided by those processes or states. Beginning with metaphysical (dis)continuism, the following section considers two potential differences, one pertaining to the objects of memories and future thoughts, the other pertaining to their reference. Epistemological (dis)continuism is discussed in section 3.

## **2 Metaphysical (dis)continuism**

### **2.1 The objects of mental time travel**

The question of *the objects of MTT* is a generalization of the traditional question of *the objects of memory* (Fernández, 2017), which concerns the nature of the entities to which we are related, in the first instance, when we remember. The two traditional answers to this question go back to the early modern period, with *direct realists*, the ancestors of current *relationalists*, arguing that, when one remembers, one is directly related to a past event itself (Reid, 2000/1764) and *indirect realists*, the ancestors of *representationalists*, arguing that, when one remembers, one is directly related to a mental representation of a past event and thus only indirectly related to the event (Locke, 1975/1689; Hume, 2011/1738). An understanding of memory as a form of MTT leads naturally to the more general question of

the nature of the entities to which we are related, in the first instance, when we mentally travel in time. And this more general question is naturally approached via an attempt to generalize the traditional answers to the traditional question. Thus one might adopt a relationalist approach and argue that, when one engages in MTT, the immediate object of one's memory or future thought is a past or future event itself, or one might adopt a representationalist approach and argue that, when one engages in MTT, the immediate object of one's memory or future thought is a representation of a past or future event.

Relationalism has a good deal of intuitive plausibility, especially with respect to memory—we ordinarily say, after all, that we remember *events*, not *representations* of events. But relationalism also faces daunting problems, especially with respect to future thought. It is difficult to see how one might, when one engages in remembering, be in direct contact with a past event. It is far more difficult to see how one might, when one engages in future thinking, be in direct contact with a future event. Past events are past, and their temporal distance from the remembering subject is a serious difficulty for the relationalist approach, but at least they once existed. Future events, in contrast, are merely possible: they do not exist at the time at which they are imagined, and they might never exist. Thus, unless relationalists are willing to commit to the extremely strong ontological claim that events exist regardless not only of their temporal location (past/future) but also of their modal location (actual/possible),<sup>4</sup> they will need to explain not only how memory can put the subject in direct contact with no-longer-existent past events but also how future thought can put the subject in direct contact with not-yet-existent future events.

Relationalists, such as Debus (2008), have responded to this problem by combining the causal theory of memory with a form of *disjunctivism* about the objects of MTT inspired

---

<sup>4</sup> See Bernecker (2008) for an argument for eternalism, the view that past events continue to exist; Bernecker does not discuss future events.

by disjunctivism about perception (see Byrne & Logue, 2009). The core idea of Debus's strategy is to argue, first, that the causal connection that, according to the causal theory, necessarily links a memory of an event to the remembered event enables the latter to serve as the direct object of the former, second, that, because no such causal connection exists in the case of future thought, the imagined event cannot be the direct object of a thought of an event, and third, that there is therefore a difference in kind, within the category of MTT, between memories and future thoughts. This emphatically discontinuist strategy, in short, consists in endorsing relationalism with respect to memory but rejecting it with respect to future thought.

While committed relationalists may find Debus's strategy appealing, Sant'Anna and Michaelian (forthcoming) have argued that disjunctivism raises new problems for relationalism by making memory into an atypical member of the broader category of MTT, thereby implying that the centrality normally assigned to memory in investigations of MTT is unjustified. There are two points to note here. First, in cases of unsuccessful memory, such as confabulation (in which one "makes up" a past event that did not occur; see Schnider, 2018), no past event is available to serve as the object of the memory. Second, the category of MTT includes not only episodic future thought but also episodic counterfactual thought (in which one imagines an event that might have but did not occur; see De Brigard, 2014a), and, in episodic counterfactual thought as well, no event is available to serve as the object of the thought.<sup>5</sup> Given these two points, the "not memory" disjunct would need to include not only future thought but also counterfactual thought and various forms of unsuccessful memory and would therefore end up dwarfing the "memory" disjunct. From a disjunctivist point of view, in other words, episodic memory ends up looking like an exception, rather than the rule.

---

<sup>5</sup> Episodic counterfactual thought is a distinct form of MTT, but we will largely refrain from discussing it here; a fuller treatment of the continuism-discontinuism debate would take it into account. We briefly discuss confabulation in section 4.

Relationalism has been relatively unpopular among philosophers of memory, and representationalism can fairly be said to be the standard approach to the objects of memory and, by extension, to the objects of MTT.<sup>6</sup> The key virtue of representationalism is that, simply because the existence of a mental representation of an event does not require the existence of the represented event, it allows us to say that all forms of MTT—memory, both successful and unsuccessful, and future and counterfactual thought—have objects of the same kind: regardless of whether one remembers an actual past event, imagines a counterfactual past event, or imagines a future event, one is related, in the first instance, to a representation, not to an event. Since we are, given standard views in the philosophy of mind, required to posit mental representations for independent reasons, moreover, representationalism, unlike relationalism, comes at no significant ontological cost.

Representationalism rules out one potential metaphysical difference between memory and future thought and is thus compatible with continuism (since, as we will see in the following section, it does not rule out other potential differences, it does not *entail* continuism). But representationalism, like relationalism, faces important problems. Building on an analogous argument in the philosophy of perception (Travis, 2004), Sant’Anna and Michaelian (forthcoming) have argued that one problem is that the representations produced by MTT are “silent” in the sense that they do not establish their own accuracy conditions. Suppose, for example, that one entertains a thought of a visit to Ottawa. The suggestion is that there is nothing internal to the thought that establishes whether it pertains to an actual past visit or, say, to a possible future visit. Accuracy conditions for an episodic thought, Sant’Anna and Michaelian argue, are established only when the thought is conjoined with further dispositions; a tendency to plan for the visit, for example, might locate the event in

---

<sup>6</sup> Because it is the standard approach, it is often simply taken for granted; see Michaelian (2016c) for one recent explicit argument for the approach.



the future. And if episodic thoughts do not establish their own accuracy conditions, it is unclear whether they qualify as genuine representations.

Another problem for representationalism is simply that it has difficulty accommodating the intuition cited at the beginning of this section, namely, that, when we remember, we remember events, not representations. Sant'Anna (forthcoming) has argued that, in light of this problem, we should consider the prospects for a hybrid view that combines aspects of representationalism with aspects of relationalism. *Hybridism* about memory, modelled on a similar view of perception (Schellenberg, 2010), is the view that the objects of memory are representations of events but that those representations are themselves constitutively determined by their causal connection to the events that they represent; this is a compromise view that asks us to reconceive the nature of memory representations by taking them to be inherently relational. Applied to MTT, hybridism supports a form of continuism that may be sufficiently moderate to appeal to discontinuists, since it implies that the objects of both memory and future thought are representations of events but that only in the case of memory are those representations determined by causal connections to events. It remains to be seen whether this moderate form of continuism constitutes a stable middle ground between more standard continuist and discontinuist positions.

## **2.2 The reference of episodic thought**

While hybridism is a promising view, representationalism remains the standard approach. Representationalism, however, raises the difficult question of the mechanism by virtue of which a memory image refers to a particular past event. This question raises another: might the mechanism by virtue of which an image refers to a particular event be at work not only in episodic memory but also in episodic future thought? The relationship that a memory bears to the event to which it refers is roughly the same as that borne by a name to the individual to

which it refers. Philosophers have developed a variety of theories of the reference of proper names (see Devitt & Sterelny, 1999), and Lopes (1996) has argued that these can be extended from *names* to *images*. If they can be extended to physical images, there would seem to be no obvious barrier to extending them further in order to provide an account of the reference of the *mental* images at issue in MTT.

Consider the description theory of reference (DTR), according to which reference is determined by *content*: names are associated with definite descriptions and refer to the entities that are singled out by the descriptions with which they are associated (Russell, 1910; Frege, 1948/1892). An utterance of ‘Toronto’, for example, might refer to Toronto because the speaker associates the description ‘the city in which the CN Tower is located’ with it and because Toronto is the city in which the CN Tower is located. Since it does not invoke causal history in order to explain reference, DTR is most naturally combined with the simulation theory of memory (STM), which does not invoke causal connection in order to explain remembering. On a combined DTR/STM account, remembering would be understood as being a matter of imagining events and the memory images it produces as having contents capable in principle of singling out events from the subject’s personal past; the account would then treat a given memory image as referring to the event singled out by its content (if there is such an event), just as a name refers to the entity that is singled out by the associated description (if there is such an entity). A memory of a visit to Toronto, for example, might refer to a particular visit to Toronto because it includes a representation of the CN Tower and because one saw the CN Tower on that visit and no other.

The DTR/STM account is compatible with continuism: because it treats reference as being underwritten by content (independent of causal history), the account implies that not only the mental images produced by remembering but also those produced by future thinking may sometimes refer to particular events. This does not entail that no other qualitative

differences between these two forms of MTT will be identified, but it does rule out one important potential difference. The DTR/STM account, however, inherits a well-known problem from the theory of reference on which it is based. Being associated with an accurate description seems to be neither necessary nor sufficient for a name to refer to an entity: a speaker who mistakenly believes of Toronto that it is the capital of Canada may, under many conditions, nevertheless refer to Toronto (and not Ottawa) when he utters ‘Toronto’. Similarly, being accurate with respect to an event seems to be neither necessary nor sufficient for a memory to refer to that event: a memory of a certain visit to Toronto can arguably get virtually everything about the visit wrong and still count as a memory of that visit, rather than another to which it happens to correspond more closely.

This problem has led many philosophers of language to turn to the causal theory of reference (CTR), according to which reference is determined by *causal history*: in the simplest case, a name is introduced via an initial “baptism”, and subsequent uses of the name refer to the thing baptized because they are appropriately causally connected to this initial use of the name (Kripke, 1980). A typical current utterance of ‘Toronto’, for example, might refer to Toronto regardless of the details of the description associated with it because the town then known as “York” was baptized in 1834 with the name ‘Toronto’ and because there is an appropriate causal connection between the utterance and the baptism. Since it does not invoke content in order to explain reference, CTR is most naturally combined with the procedural causal theory of memory (PTM). The distinguishing feature of PTM is in fact its denial—motivated in part by a desire to reconcile the core of the causal theory of memory with recent contentless conceptions of memory traces (see, e.g., De Brigard, 2014b; Hutto & Peeters, 2018)—of the claim made by the classical causal theory that the causal connection constitutive of remembering involves the transmission of content from experience to retrieved memory. According to PTM, the causal connection in question links not the *content*

of the experience to the *content* of the retrieved memory but rather the constructive *process* that produces the experience to the reconstructive *process* that produces the retrieved memory: in virtue of certain brain-level similarities between the two processes (see Perrin, 2018 for details), the former may affect the fluency of the latter, thus securing a causal connection between the two despite the fact that no content is transmitted from one to the other. On a combined CTR/PTM account, a memory would be understood as referring to a given event just in case it causally derives, in this “procedural” manner, from an experience of that event. A memory of a visit to Toronto, for example, might refer to a certain visit because the reconstructive process that generates the memory is facilitated by the earlier constructive process of experiencing that visit.

The CTR/PTM account supports discontinuism: because it treats reference as being underwritten by causal history (independent of content), and because future thoughts are not caused by the events that they are about, the account implies that future thought, unlike memory, never involves reference to particular events (cf. Debus, 2014). And if memories sometimes refer to particular events and future thoughts do not, this would seem to amount to a fundamental metaphysical difference between them. Like the DTR/STM account, however, the CTR/PTM account inherits a well-known problem from the theory of reference on which it is based. Suppose that the baptism in 1834 assigned the name ‘Toronto’ to Lake Ontario but, due to miscommunication, was taken by subsequent speakers to have assigned it instead to the neighbouring city; because a typical current utterance of ‘Toronto’ is causally connected to the baptism, CTR implies that it refers to Lake Ontario, despite the fact that Toronto is the dominant causal source of the information associated by the speaker with the name. One might suspect that, because PTM denies that content is transmitted from experiences of events to memories of them, an analogous problem cannot arise for the CTR/PTM account, but the account does in fact face a version of the problem: the fluency of

the reconstructive process that generates a memory of a certain visit to Toronto, for example, might be affected not only by the earlier constructive process of experiencing that visit but also—and possibly primarily—by the processes of experiencing other, similar visits.

This problem motivates the hybrid theory of reference (HTR), according to which a name refers to the *dominant causal source* of the content associated with it (Evans, 1982): in the scenario described above, a typical utterance of ‘Toronto’ would, if HTR is right, refer to Toronto, rather than Lake Ontario, because the content that the speaker associates with it derives primarily from the city, not the lake. HTR is a hybrid theory in the sense that it invokes both causation and content and is thus combined most naturally with CTM, which requires that the causal connection between a present memory image and a past event involve the transmission of content from the experience of the event to the image; if HTR is right, this causal connection is precisely of the right sort to secure reference to the event that gave rise to the experience from which content is transmitted. On a combined HTR/CTM account, then, memories refer to events because they are causally connected to experiences of those events and because the relevant causal connection involve the transmission of content from the experiences in question.

The HTR/CTM account, like the CTR/PTM account, supports discontinuism: because it treats reference as being underwritten by the transmission of content, the account implies—given that future thoughts are never thus connected to the events that they are about—that future thought, unlike memory, never involves reference to particular events. The account, however, inherits a pair of problems from the theory of memory on which it is based. First, procedural causal theorists will object that it is straightforwardly incompatible with arguments for the contentless conception of memory traces. Second, simulation theorists will object that it cannot in fact explain why a given memory refers to a given event: assuming that we grant that remembering involves the transmission of content, research on the

reconstructive character of remembering will force us to acknowledge that content may be transmitted not only from the experience of the event that a memory is about but from experiences of other events as well and indeed that, in some cases, the event that a memory is about may not be the dominant source of the content of the memory.

Of the three accounts of the reference of memory that we have discussed, one (the DTR/STM account) favours continuism and two (the CTR/PTM and HTR/CTM accounts) favour discontinuism. None of these accounts is fully satisfactory: the DTR/STM and CTR/PTM accounts inherit problems from the relevant theories of reference, while the HTR/CTM account inherits problems from the relevant theory of memory. Some of these problems may eventually be solved, but it is unclear at present what implications an adequate account of the reference of memory will ultimately have for the continuism-discontinuism debate. It is, however, clear that the debate between metaphysical continuists and discontinuists tends to boil down to a disagreement over the necessity, for the occurrence of remembering, of a causal connection with the remembered event, with discontinuists endorsing one or another version of the causal theory and continuists rejecting it.

### **3 Epistemological (dis)continuism**

Whereas metaphysical (dis)continuism rejects (accepts) the existence of fundamental differences between episodic memory and episodic future thought, understood as cognitive processes or mental states, epistemological (dis)continuism rejects (accepts) the existence of fundamental differences between the knowledge of past events and the knowledge of future events that is provided by those processes or states. Given that the normative (including the epistemic) supervenes on the descriptive, adopting epistemological discontinuism commits one to the adoption of some form of metaphysical discontinuism. We will not discuss the

relationship between particular forms of epistemological discontinuism and particular forms of metaphysical discontinuism in any detail here, but we will show that the debate between epistemological discontinuism and continuism, like that between metaphysical discontinuism and continuism, tends to boil down to a disagreement over causation. This section considers three potential epistemic discontinuities, one concerning the epistemic openness of the future, another concerning the directness of our knowledge of the past, and a third concerning immunity to error through misidentification in episodic memory and episodic future thought.

### **3.1 The epistemic openness of the future**

One motivation for epistemological discontinuism is the natural thought that, even if future thought can provide us with knowledge of future events, that knowledge is bound to be far less secure than is the knowledge of past events with which we are provided by memory, for, if there is controversy over whether the future is *metaphysically* open (Borghini & Torrenco, 2013), there is no controversy over whether it is *epistemically* open: even if there is a fact of the matter about what will happen in the future, we cannot be *certain* about what will happen. Indeed, even if there is a fact of the matter about what will happen, we cannot be certain that *anything at all* will happen. It may be improbable, but it is perfectly consistent with the evidence available to us that the entire world will blink out of existence five minutes from now. The past, in contrast, is not epistemically open to nearly the same degree. We may be able to doubt the details of many of our memories, but we can at least be certain that the world did not blink into existence five minutes ago. There would thus seem to be a qualitative difference between the knowledge provided by memory and the knowledge provided by future thought.

Natural though this thought may be, it is mistaken. As Russell (2005/1921) pointed out, it is in fact perfectly consistent with the evidence available to us that the world blinked

into existence five minutes ago, complete with our memories, no matter how detailed or subjectively convincing these might be. This is just a particular instance of a general sceptical point in epistemology: if we set the standards for knowledge sufficiently high, requiring not just reliability but certainty, we thereby deprive ourselves of virtually all knowledge, including knowledge of the very existence of the external world—if certainty is a prerequisite for knowledge, we can have knowledge neither of future events, nor of past events, nor even of present events. Thus, if we are to avoid scepticism, the standards for knowledge should be set well below the level of certainty. And if knowledge presupposes only some lower level of reliability, the openness of the future need no longer give us any reason to suppose that there is a qualitative epistemological difference between memory and future thought, for the past is to some extent open as well: Michaelian (2016a) thus argues that, though our beliefs about future events are likely to be significantly less reliable than our beliefs about past events, this may nevertheless amount to a merely quantitative difference—a difference of degree, rather than a difference in kind—between our knowledge of future events and our knowledge of past events.

### **3.2 The directness of our knowledge of the past**

Another motivation for epistemological discontinuism is the thought that, though memory may be on a par with future thought with respect to *indirect* or inferential knowledge, there is a qualitative difference between them with respect to their capacity to provide us with *direct* or noninferential knowledge. Kneale (1971), for example, while granting that both memory and future thought are imperfectly reliable, nevertheless argues that it is part of the very concept of memory that a memory “should have as a part-cause the occurrence of the event recollected” (2). This is, in effect, to invoke the causal theory of memory. But while the claim that memory but not future thought requires a causal connection with the event in question



entails a metaphysical difference between the two, it does not entail an epistemological difference between them unless conjoined with the further claim that knowledge of an event (as opposed to mere true belief about it) requires a causal connection with it. Thus Kneale's view presupposes a *causal theory of knowledge* (Goldman, 1967).

Causal theories of knowledge rule out not only knowledge of future events but also a variety of other kinds of knowledge (e.g., of mathematical facts). They thus have few adherents and have largely been supplanted by *reliabilist theories of knowledge* according to which knowledge requires reliability but not necessarily a causal connection (Goldman, 2012). Given reliabilism, a representation produced by a process that involves a causal connection to the represented event might well qualify as knowledge, but, if it so qualifies, it does so in virtue of the reliability of the process that produced it. A reliable process that does not involve a causal connection to the events of which it produces representations is just as capable of producing representations that qualify as knowledge. Michaelian (2016a) thus argues that the "directness" of memory knowledge (in Kneale's causal sense) is irrelevant: if reliabilism is right, then, if memory and future thought are both sufficiently reliable, they are both capable of providing us with knowledge of events.

### **3.3 Immunity to error through misidentification in episodic memory and episodic future thought**

Neither the epistemic openness of the future nor the directness of our knowledge of the past, then, seems to give us good reason to endorse epistemological discontinuism. A third and final potential epistemological discontinuity between episodic memory and episodic future thought pertains to their *immunity to error through misidentification* (IEM) or lack thereof. There are two senses in which memory might be thought to be immune to error through

misidentification. First, it might be *factually* IEM.<sup>7</sup> If it is, then, given the way memory actually works, if, for example, I have a memory on the basis of which I believe that I took a walk in Ottawa last year, then I might be wrong *that* a walk was taken, but, if I am not wrong about that, I cannot be wrong in thinking that it was *I* who took the walk. Second, it might be *logically* IEM. If it is, then, *no matter how memory might work*, if I have a memory on the basis of which I believe that I took a walk in Ottawa last year, then, if I am not wrong about whether a walk was taken, I cannot be wrong in thinking that it was *I* who took the walk (McDowell, 1997; Hamilton, 2007). It is plausible that memory is factually IEM, but factual IEM does not imply logical IEM,<sup>8</sup> and Perrin (2016) has argued that this points to a fundamental difference between memory and future thought: memory may be factually IEM, but only future thought is logically IEM.

Perrin's argument appeals, first, to the claim that the sort of causal connection invoked by the causal theory of memory plays a role in determining the identities of the subjects—other than the subject whose experience is remembered—who figure in the memory. Consider, for example, a scenario in which I have a memory on the basis of which I believe that I took a walk with Paul but in which, unbeknownst to me, I in fact took a walk with Peter, Paul's identical twin brother. In this scenario, in virtue of the causal connection between the memory of the walk and the experience of it, I remember Peter but misidentify him as Paul. Memory, in other words, is not even factually IEM with respect to the identities of subjects other than the subject whose experience is remembered. His argument appeals, second, to the claim that the same causal connection determines the identity of the subject whose experience is remembered. Given the way memory actually works, this causal

---

<sup>7</sup> We will use "IEM" as an abbreviation for both "*immune* to error through misidentification" and "*immunity* to error through misidentification".

<sup>8</sup> Note that, given how we have defined factual IEM, logical IEM implies factual IEM. If we were instead to define factual IEM as *mere* factual IEM, then factual IEM and logical IEM would be mutually exclusive.

connection only ever links the subject at the time of remembering to the very same subject at the time of experiencing. Hence *memory is factually IEM* with respect to the identity of the subject whose experience is remembered. But if memory were to work differently, the causal connection might link the subject at the time of remembering to a different subject at the time of experiencing. Suppose, for instance, that a “memory transplant” technology were to be invented.<sup>9</sup> It could then turn out that, when I have a memory on the basis of which I believe that I took a walk, I am not wrong in thinking that a walk was taken but am wrong in thinking that it was me who took the walk, for it might be that I have retrieved a transplanted memory. In this case, I might misidentify another subject as myself: *memory is not logically IEM* with respect to the identity of the subject whose experience is remembered (see Coliva, 2006). Perrin’s argument appeals, finally, to the claim that, in contrast to memory, *future thought is logically IEM*. Consider a scenario in which I imagine taking a walk next year with Paul. Since no causal connection is available to determine the identity either of the subject with whom the walk will be taken or the identity of the subject who will take the walk, both identities are in effect stipulated by the imagining subject—that is, they are determined by his intentions (see Recanati, 2007). Thus, regardless of whether I intend to imagine myself taking the walk or to imagine another subject taking the walk, there is no possibility of error with respect to whose experience I am imagining.

A difference between memory and future thought with respect to logical IEM would mean that there is an important kind of false belief to which we are susceptible in one form of MTT but not the other and would thus amount to a fundamental epistemological discontinuity between them. But alternative continuist treatments of the question of IEM in episodic memory and episodic future thought are available. Michaelian (2016a) defends a view on which both memory and future thought are sometimes logically and therefore factually IEM

---

<sup>9</sup> See Shoemaker (1970) and Parfit (1984) on “quasi-memory”.

and sometimes factually but not logically IEM. The basic thought behind the view is that, once we admit that memory is a form of MTT, then we have to admit that any mechanism capable of determining identity in future thought might sometimes also be at work in memory, and vice versa. Thus, while Perrin maintains that identity is determined by causation in future thought and by stipulation in memory, Michaelian argues, first, that identity is *sometimes* determined by stipulation not only in future thought but also in memory (and hence that memory, like future thought, is sometimes logically—and therefore factually—IEM) and, second, that identity is *sometimes* determined by causation not only in memory but also in future thought (and hence that future thought, like memory, is sometimes factually but not logically IEM).

On the one hand, the suggestion that future thoughts are sometimes logically IEM derives its plausibility from the fact that there is no causal connection between future events and present representations thereof, leaving the subject's intentions to determine identity, but, if memory is just as much a form of MTT as is future thought, then there is no guarantee that, when a subject remembers, there will inevitably be a causal connection between the past event and his present representation thereof, in which case it is just as plausible that identity is determined by the subject's intentions. If it is, then the memory is logically and hence factually IEM. On the other hand, because the contents of simulations of both past and future events derive in part from the subject's experiences of past events, the identity of the subject whose experience is represented can be inherited via a causal connection between the present representation and a past experience not only in memory but also in future thought. Consider a simple case in which I, in a first step, remember taking a walk in Ottawa last year and, in a second step, imagine taking a walk in Ottawa next year by mentally projecting that experience forward in time. If one claims that the causal connection confers factual IEM on the memory, there would seem to be little reason to deny that it confers factual IEM on the

future thought. By the same token, if one takes the in-principle possibility of memory transplants to rule out logical IEM with respect to the memory, one should likewise take it to rule out logical IEM with respect to the future thought: if a memory transplant were to lead to error due to misidentification at the first step, that error would presumably be inherited at the second step—I would be imagining not my own future walk but rather that of the subject whose memory has been transplanted.

If these arguments are right, then there will be cases of both future thought and memory that are logically and therefore factually IEM, and there will be cases of both memory and future thought that are factually but not logically IEM. Discontinuist replies to the arguments are, of course, available. Regarding the second part of the argument, the discontinuist might insist that memory is involved in the process of imagining and that, if the subject instead imagines himself taking a walk next year without drawing on a memory of himself taking a walk last year, the identity of the relevant subject will be determined entirely by his intentions. Regarding the first part of the argument, the discontinuist might insist that, if there is no causal connection to the “remembered” event, then the subject is not really remembering but rather imagining. If these replies are successful, then we are back to discontinuism: future thought is logically and therefore factually IEM, and memory is factually but not logically IEM.

Another continuist alternative is suggested by Fernández’s (2014) argument for the claim that memory is logically IEM. Fernández points out that the phenomenology of episodic memories is perspectival. The perspectival character of memories arguably derives from that of the past perceptual experiences on which they are based. On Fernández’s view, the perspectival character of perception is due to the fact that it represents relations between the perceiving subject and objects in the external world. For example, as I see the CN Tower in the distance, I represent myself as occupying a certain position relative to it. Consequently,

when I later remember having seen the CN Tower, I represent a determinate subject, myself, as having occupied that position. I thus necessarily represent *myself* as having had the experience. To be sure, I can be mistaken about whether I actually had *the experience*, but I cannot be mistaken about whether I am *the subject* represented as having had the experience. Crucially, this is so regardless of what happened in the past and, in particular, regardless of whether my current memory results from a memory transplant.

If Fernández's argument is successful, then memory is logically IEM. If this argument were to be combined with Perrin's argument for the claim that future thought is logically IEM, we would in effect have an argument for a combined view on which *both* memory and future thought are logically IEM. This combined view, however, sees entirely different mechanisms as being responsible for IEM in memory and in future thought and is thus of doubtful coherence. Moreover, a discontinuist might argue, first, that the perspectival character of memory does not fix the identity of the subject whose experience is being remembered but only the fact that the experience involved perception from a certain perspective. Memory representations would thus be silent with respect to identity, in which case the possibility of memory transplants would again entail the logical possibility of error due to misidentification. The discontinuist might argue, second, that Fernández assumes that, when one remembers, one represents the experience and the experiencing subject, whereas it is more parsimonious to assume that only the relevant event is explicitly represented (Zahavi, 2003). If the experiencing subject does not figure in the representation of the event, we arrive, again, at the possibility of memory transplants, entailing the logical possibility of error due to misidentification.

The continuist might be able to successfully defend a variety of epistemological continuism based on Fernández's claim that memory is logically IEM, but a variety that sees both memory and future thought as sometimes logically and therefore factually IEM and

sometimes factually but not logically IEM is more likely to be coherent. An advocate of a continuist position of this sort might argue that the presence of an appropriate causal connection in cases of memory and the absence of such a connection in cases of imagination has, in the discontinuist objections to the position outlined above, the status of a dogma: if appropriate causation is indeed the dividing line between remembering and imagining, then there will, indeed, be a fundamental metaphysical difference between those two processes, in which case we should not be surprised to find fundamental epistemic differences between them as well, but, since the continuism-discontinuism debate was triggered in the first place by the fact that MTT research calls the role of appropriate causation in distinguishing between memory and imagination into question, the discontinuist is hardly entitled to take this role for granted. How, then, are we to settle the matter of the role of appropriate causation in distinguishing between memory and imagination? We noted at the outset that the available empirical evidence does not suffice to decide between continuism and discontinuism. It may, however, favour either the causal theory or the simulation theory. Since the former tends to support discontinuism and the latter to support continuism, the resolution of the continuism-discontinuism debate may ultimately be decided by the resolution of the causalist-simulationist debate.

#### **4 The future of the continuism-discontinuism debate: Future-oriented confabulation?**

The continuism-discontinuism debate is multifaceted, and there are a number of other issues that this chapter might have broached (see Perrin & Michaelian, 2017; Sant'Anna, 2018; for a recent review from a psychological perspective, see Addis, 2018). Before concluding, we briefly explore just one of these: the implications of continuism and discontinuism for the possibility of future-oriented confabulation. *Confabulation* is, very roughly, a form of

unsuccessful memory in which subjects unable to remember “make up” more or less plausible past events by, for example, combining aspects of different experienced events or radically displacing events in time (see Schnider, 2018). In the present context, it is natural to inquire into the utility of the concept of *future-oriented* confabulation analogous to this familiar concept of (past-oriented) confabulation. Whether we can make sense of the concept of future-oriented confabulation depends, once again, on our stance with respect to the necessity of a causal connection for the occurrence of memory.

There are accounts of the relationship between confabulation and remembering based both on the causal theory and on the simulation theory. *Causalist* accounts (Robins, 2016, forthcoming a, b; Bernecker, 2017) treat the existence of an appropriate causal connection with the apparently remembered event as making the difference between remembering and confabulating. *Simulationist* accounts (Michaelian 2016b, forthcoming), in contrast, see the difference as being a matter of reliability, in effect characterizing confabulation as unreliable imagination. If what is distinctive of past-oriented confabulation is, as the causal theorist would have it, lack of causal connection, then the notion of future-oriented confabulation will make little sense, simply because future thought never involves a causal connection to the represented event. If what is distinctive of past-oriented confabulation is, instead, as the simulation theorist would have it, lack of reliability, then the notion of future-oriented confabulation may well make sense.

It is, of course, one thing for the concept to make sense and quite another for it to correspond to an empirical phenomenon, and one might maintain that, since successful future thoughts, unlike successful memories, need not correspond to actual events, there is simply no need for a concept of future-oriented confabulation. There is, however, considerable work showing that confabulators display not only defective remembering but also defective future thinking, in the sense that they produce representations of future events that fail to correspond



to events that they are likely to experience (see, e.g., Dalla Barba et al., 1997; Schnider, 2018), suggesting that there is indeed a need for a concept of future-oriented confabulation. Whether the concept of future-oriented confabulation is viable will ultimately depend on whether a notion of reliability applicable to episodic future thinking (and potentially to episodic imagining more broadly) can be worked out, and this remains to be done.

Our aim in this chapter has been to review arguments for and against continuism and discontinuism; we have taken a stand in favour of neither view. The primary message of the chapter is thus that it is likely that the resolution of the continuism-discontinuism debate will depend on a resolution of the question of the necessity of a causal connection for the occurrence of memory. The causal theory of memory has been enormously influential among philosophers of memory (Michaelian & Robins, 2018), and, from the perspective of the philosophy of memory, the burden of proof tends to fall on those who would endorse the simulation theory. Things may, however, look very different from the perspective of the philosophy of imagination. We therefore close by inviting philosophers of imagination to consider what implications their accounts of the nature of episodic imagination might have for the continuism-discontinuism debate.

**Acknowledgements:** Thanks to Anna Abraham for detailed comments on an earlier draft of this chapter. The chapter is based in part on work done during a fellowship awarded to Kourken Michaelian by the Fondation Maison des sciences de l’homme.

## References

- Addis, D. R. (2018). Are episodic memories special? On the sameness of remembered and imagined event simulation. *Journal of the Royal Society of New Zealand*, 48(2-3), 64-88.
- Arcangeli, M. (Forthcoming). *Supposition and the Imaginative Realm*. London: Routledge.

- Bernecker, S. (2008). *The Metaphysics of Memory*. Springer.
- Bernecker, S. (2017). A Causal Theory of Mnemonic Confabulation. *Frontiers in Psychology*, 8, 1207.
- Bernecker, S., & Michaelian, K. (Eds.). (2017). *The Routledge Handbook of Philosophy of Memory*. London: Routledge.
- Borghini, A., & Torrenco, G. (2013). The metaphysics of the thin red line. In F. Correia & A. Iacona (Eds.), *Around the Tree* (pp. 105-125). Springer.
- Brewer, W. F. (1996) What is recollective memory? In D. C. Rubin (Eds.), *Remembering Our Past: Studies in Autobiographical Memory* (pp. 19-66). Cambridge: Cambridge University Press.
- Byrne, A., & Logue, H. (Eds.). (2009). *Disjunctivism: Contemporary Readings*. Cambridge: MIT Press.
- Coliva, A. (2006). Error through misidentification: Some varieties. *The Journal of Philosophy*, 103(8), 403-425.
- Dalla Barba, G., Cappelletti, J. Y., Signorini, M., & Denes, G. (1997). Confabulation: Remembering 'another' past, planning 'another' future. *Neurocase*, 3(6), 425-436.
- Debus, D. (2008). Experiencing the past: A relational account of recollective memory. *Dialectica*, 62(4), 405-432.
- Debus, D. (2014). 'Mental time travel': Remembering the past, imagining the future, and the particularity of events. *Review of Philosophy and Psychology*, 5, 333–350.
- De Brigard, F. (2014a). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese*, 191(2), 155-185.
- De Brigard, F. (2014b). The nature of memory traces. *Philosophy Compass*, 9(6), 402-414.
- Devitt, M., & Sterelny, K. (1999). *Language and Reality: An Introduction to the Philosophy of Language*. Cambridge: MIT Press.
- Evans, J. (1982). *The Varieties of Reference*. Oxford: Clarendon Press.
- Fernández, J. (2014). Memory and immunity to error through misidentification. *Review of Philosophy and Psychology*, 5(3), 373-390.
- Fernández, J. (2017). The intentional objects of memory. In S. Bernecker & K. Michaelian (Eds.), *The Routledge Handbook of Philosophy of Memory* (pp. 88-99). London: Routledge.
- Frege, G. (1948/1892). Sense and reference. *The Philosophical Review*, 57(3), 209-230.
- Frise, M. (2015). Epistemology of Memory. In J. Fieser & B. Dowden (Eds.), *Internet Encyclopedia of Philosophy*. Retrieved from: <http://www.iep.utm.edu/epis-mem/>

- Gérardin-Laverge, L. (2017). Mémoire constructive, imagination et voyage mental dans le temps. *Cahiers philosophiques*, (2), 23-40.
- Goldman, A. I. (1967). A causal theory of knowing. *The Journal of Philosophy*, 64(12), 357-372.
- Goldman, A. I. (2012). *Reliabilism and Contemporary Epistemology: Essays*. Oxford: Oxford University Press.
- Hamilton, A. (2007). Memory and self-consciousness: immunity to error through misidentification. *Synthese*, 171, 409-417.
- Hume, D. (2011/1738). *A Treatise of Human Nature*. Oxford: Clarendon Press.
- Hutto, D. D. & Peeters, A. (2018). The roots of remembering: Radically enactive recollecting. In K. Michaelian, D. Debus & D. Perrin (Eds.), *New Directions in the Philosophy of Memory* (pp. 97-118). London: Routledge.
- Kind, A., & Kung, P. (Eds.). (2016). *Knowledge Through Imagination*. Oxford: Oxford University Press.
- Kneale, M. (1971). Our knowledge of the past and of the future. *Proceedings of the Aristotelian Society*, 72:1-12.
- Kripke, S. (1980). *Naming and Necessity*. Cambridge: Harvard University Press.
- Locke, J. (1975/1689). An Essay Concerning Human Understanding. In P. Nidditch (Ed.), *The Clarendon Edition of the Works of John Locke: An Essay Concerning Human Understanding*. Oxford: Oxford University Press.
- Lopes, D. (1996). *Understanding Pictures*. Oxford: Clarendon Press.
- Martin, C. B., & Deutscher, M. (1966). Remembering. *The Philosophical Review*, 75(2), 161-196.
- McDowell, J. (1997). Reductionism and the first person. In J. Dancy (Eds.), *Reading Parfit* (pp. 230-250). Oxford: Blackwell.
- Macpherson, F., & Dorsch, F. (Eds.). (2018). *Perceptual Imagination and Perceptual Memory*. Oxford: Oxford University Press.
- Michaelian, K. (2016a). Against discontinuism: Mental time travel and our knowledge of past and future events. In K. Michaelian, S. B. Klein, & K. K. Szpunar (Eds.), *Seeing the Future: Theoretical Perspectives on Future-Oriented Mental Time Travel* (pp. 69-92). Oxford: Oxford University Press.
- Michaelian, K. (2016b). Confabulating, misremembering, relearning: The simulation theory of memory and unsuccessful remembering. *Frontiers in Psychology*, 7, 1857.

- Michaelian, K. (2016c). *Mental Time Travel: Episodic Memory and Our Knowledge of the Personal Past*. Cambridge: MIT Press.
- Michaelian, K. (Forthcoming). Confabulating as unreliable imagining: In defence of the simulationist account of unsuccessful remembering. *Topoi*.
- Michaelian, K., Klein, S. B., & Szpunar, K. K. (Eds.). (2016). *Seeing the Future: Theoretical Perspectives on Future-Oriented Mental Time Travel*. Oxford: Oxford University Press.
- Michaelian, K., & Robins, S. (2018). Beyond the causal theory? Fifty years after Martin and Deutscher. In K. Michaelian, D. Debus & D. Perrin (Eds.), *New Directions in the Philosophy of Memory* (pp. 12-32). London: Routledge.
- Nanay, B. (2015). Perceptual content and the content of mental imagery. *Philosophical Studies*, 172(7), 1723-1736.
- Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.
- Perrin, D. (2016). Asymmetries in Subjective Time. In K. Michaelian, S. B. Klein, & K. K. Szpunar (Eds.), *Seeing the Future: Theoretical Perspectives on Future-Oriented Mental Time Travel* (pp. 39-61). Oxford: Oxford University Press.
- Perrin, D. (2018). A case for procedural causality in episodic recollection. In K. Michaelian, D. Debus & D. Perrin (Eds.), *New Directions in the Philosophy of Memory* (pp. 33-51). London: Routledge.
- Perrin, D. & Michaelian, K. (2017). Memory as mental time travel. In S. Bernecker and K. Michaelian (Eds.), *The Routledge Handbook of Philosophy of Memory* (pp. 228-239). London: Routledge.
- Recanati, F. (2007). *Perspectival Thought: A Plea for (Moderate) Relativism*. Oxford: Oxford University Press.
- Reid, T. (2000/1764). *An Inquiry into the Human Mind on the Principles of Common Sense*. Pennsylvania State University Press.
- Robins, S. K. (2016). Misremembering. *Philosophical Psychology*, 29(3), 432-447.
- Robins, S. K. (Forthcoming a). Confabulation and constructive memory. *Synthese*.
- Robins, S. K. (Forthcoming b). Mnemonic confabulation. *Topoi*.
- Russell, B. (1910). Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society* 11, 108-128.
- Russell, B. (2005/1921). *Analysis of Mind*. London: Routledge.
- Sant'Anna, A. (2018). Mental time travel and the philosophy of memory. *Unisinos Journal of Philosophy* 19(1): 52-62.

- Sant'Anna, A. (Forthcoming). The hybrid contents of memory. *Synthese*.
- Sant'Anna, A. & Michaelian, K. (Forthcoming). Thinking about events: A pragmatic account of the objects of episodic hypothetical thought. *Review of Philosophy and Psychology*.
- Schellenberg, S. (2010). The particularity and phenomenology of perceptual experience. *Philosophical Studies*, 149(1), 19-48.
- Schnider, A. (2018). *The Confabulating Mind: How the Brain Creates Reality*. Oxford: Oxford University Press.
- Shoemaker, S. (1970). Persons and their pasts. *American Philosophical Quarterly*, 7(4), 269-285.
- Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and brain sciences*, 30(3), 299-313.
- Travis, C. (2004). The silence of the senses. *Mind*, 113(449), 57-94.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory* (pp. 381-402). New York: Academic Press.
- Tulving, E. (1985). *Elements of Episodic Memory*. Oxford: Oxford University Press.
- Werning, M., & Cheng, S. (2017). Taxonomy and unity of memory. In S. Bernecker & K. Michaelian (Eds.), *The Routledge Handbook of Philosophy of Memory* (pp. 7-20). London: Routledge.
- Zahavi, D. (2003). *Husserl's Phenomenology*. Stanford: Stanford University Press.