A dilemma for simulationism*

Simulationism is a leading philosophical theory of remembering characterized by three ideas: continuism (remembering and imagining are constituted by one and the same natural kind), anticausalism (one may remember an event without an appropriate causal connection to it), and mnemic reliabilism (successful remembering requires process reliability). At the heart of this radical trio of claims is *identity-of-process:* a single simulative process *just is* the psychological basis of both remembering and imagining. Though simulationism has faced much criticism, this has seldom engaged with the theory on its own methodological terms. In this paper, we argue that commitment to identity-of-process generates a dilemma. If simulationists take identity-of-process seriously, they eliminate remembering as a natural kind, undermining the very explanatory ambitions of a naturalistic theory of remembering, and lose their grip on mnemic success and error. If they abandon identity-of-process, however, simulationism becomes a discontinuist view, one compatible with causalism, thus jeopardizing its status as a distinct or radical proposal. We conclude by examining strategies for escaping the dilemma and reflect on its implications for broader debates concerning (dis)continuism and the individuation of mental kinds, (anti)causalism, and naturalistic methodology in philosophy of memory.

1. Introduction

Simulationism is a leading philosophical theory of remembering (Michaelian 2016a, 2024; De Brigard 2014, 2023). It is grounded in a scientific research program that individuates mental kinds in terms of the neurocognitive systems and component processes that underlie them. Some simulationists have endorsed an *identity-of-process* view, according to which a single process underlies multiple intuitively distinct activities, including remembering and imagining. The process is an operation of a system whose function has been characterized as mental time travel (Tulving 2005; Suddendorf & Corballis 2007), constructive episodic simulation (Schacter & Addis 2007; Addis 2020) and scene construction (Hassabis & Maguire 2009). On this view, remembering occurs when the process is recruited for the purpose of representing an event from the organism's past.

Simulationists have argued for three important theses. First, *continuism*: remembering and imagining are activities of the same natural kind. According to continuism, while some differences

* The authors contributed equally.

between remembering and imagining may exist, these differences are of degree not kind. Second, *anticausalism*: remembering does not require any ('appropriate') causal connection to the remembered event. Just as we can imagine a future event to which we are not causally connected, we can likewise remember a past event relying solely on information causally connected to sources other than our own past experience of it. Third, *mnemic reliabilism*: the reliable functioning of the simulative process is necessary for successful remembering. We can account for different memory errors (e.g., misremembering or confabulation) by characterizing the ways in which the simulative process, and associated processes, malfunction.

These theses are radical and counterintuitive, and have drawn significant attention and criticism from philosophers of memory. *Discontinuists* argue for important differences between remembering and imagining—phenomenological, attitudinal, and linguistic (Perrin 2016; Robins 2020; Liefke & Werning 2024). *Causalists* challenge the claim that we can genuinely remember an event to which we are not (in the appropriate sense) causally connected (Perrin 2021; Moran 2022; Langland-Hassan 2022). And philosophers with otherwise divergent commitments have objected to mnemic reliabilism (Bernecker 2017; Werning 2020; Lai 2024). While critics have offered valuable insights, they have often failed to fully appreciate the nature of the simulationist program and the source of its counterintuitive claims. Consequently, as Camillo (2025) has recently argued, these objections have rarely posed a significant challenge to simulationism.

Indeed, simulationism's methodological commitments, and the relationships among its various theses, have gone largely unexplored. This is particularly significant because it is not *prima facie* obvious that simulationism constitutes a methodologically coherent view. To pull on the most obvious thread: if remembering and imagining form a single natural kind, then there is no type of state that aligns with our intuitive concept of remembering and whose causal profile forms the basis for interesting generalizations. If there are no interesting generalizations about remembering *per se*, it is unclear what could underwrite the simulationist's talk of the accuracy/success conditions of remembering (and tabling of errors) as distinct from those of imagining.

In this paper, we examine simulationism through the lens of these methodological tensions and pose a central dilemma for its proponents. If simulationism endorses identity-of-process, it eliminates remembering as a natural kind, and thus as a target of productive theorizing. Moreover, the version of simulationism that is committed to identity-of-process provides no adequate basis for either a substantive anticausalism, or an account of successful remembering grounded in the reliability of a

specifically mnemic process. If, on the other hand, simulationism abandons identity-of-process, it becomes a discontinuist view compatible with various forms of causalism, jeopardizing its status as a radical or genuinely distinct option in the theoretical landscape.

Our analysis focuses primarily on Michaelian's (2016a, 2024) account, which remains the most systematic and influential articulation of simulationist ideas. In the course of our analysis, however, we bring this account into dialogue with other simulationist proposals in both philosophy and psychology, tracing where their commitments align and diverge. The paper is structured as follows. In Section 2, we provide an overview of simulationism, emphasizing its methodological basis in a particular form of philosophical naturalism. Section 3 develops the central dilemma, bringing out the tension between identity-of-process and the broader commitments of the simulationist program. In Section 4, we consider potential strategies for escaping the dilemma, focusing on amendments that would require either methodological or theoretical revisions to simulationism. Section 5 concludes.

2. Naturalism and simulationism: Toward identity-of-process

In 2.1, we examine the form of philosophical naturalism that underwrites the simulationist program. In 2.2, we introduce simulationism in more detail, focusing on the empirical and theoretical motivations for endorsing the identity-of-process thesis, elucidated in 2.3.

2.1. Philosophical naturalism

Simulationism rests on a complex methodological framework that warrants a variety of labels. Here we adopt the label *naturalism*, favored by the leading proponent of simulationism (Michaelian 2016a, 2016b, 2024; Andonovski & Michaelian 2024).[1] We do so for dialectical convenience but with some reservations. There are two principal reasons for having such reservations. First, the term has been employed in many ways, both by self-styled naturalists and their opponents (e.g., see the essays in Clark et al. 2015 and Kemp et al. 2024). Second, the cluster of commitments that underwrite the simulationist program pertain both to the nature of investigated phenomena and to the roles science and philosophy play in such investigation, thus covering aspects of both *metaphysical* and

[1] Perhaps tellingly, De Brigard (2014, 2023, 2025) does not characterize his preferred approach as naturalistic, but rather as *inclusive* and as appealing "not only [to] introspective evidence from our phenomenology and conceptual analyses…but also [to] empirical results from the sciences of memory – particularly from cognitive psychology and neuroscience" (2023, p. 4). Commitment to such an inclusive approach, which De Brigard likens to a kind of reflective equilibrium, will likely make him less sympathetic to the eliminativist conclusions that, as we will see, theorists like Michaelian and Addis may be forced to endorse.

methodological varieties of naturalism. For these reasons, while we continue to use *naturalism* throughout the paper, the reader should be aware that our usage may be somewhat idiosyncratic. They should thus attend closely to the specific commitments articulated below. It is to these—rather than to any preconceived notion of naturalism—that our use of the term should be anchored.

In what follows, we identify three key commitments associated with naturalism. First, the world is thought to exhibit a *natural-kind* structure, consisting of natural categories of entities that are the proper targets of scientific inquiry. In a recent book, Khalidi (2023) articulates this view with a focus on entities studied in the sciences of the mind:

> In investigating mental objects, states, capacities, events, processes, and other entities, we are usually investigating types not tokens, that is, not unique particulars, but types or *kinds* of them. Specifically, we are interested in which of these types or kinds are real or "natural," or in standard philosophical parlance: *natural kinds* (p. 3)

The view that the world consists of natural kinds—genuinely natural ways of classifying entities—is widespread in contemporary analytic philosophy and underpins influential versions of both naturalism and scientific realism (see Bird & Tobin 2024; Brzović 2024). That said, there are ongoing debates about how natural kinds should be characterized, with traditional essentialist views gradually giving way—in popularity, at least—to more liberal cluster views that view kinds as co-occurring properties linked by causal or functional relations (Boyd 1991; Slater 2015; Khalidi 2018). While simulationists typically endorse such cluster views (see, e.g., Michaelian 2015), what matters for our purposes is the commitment to the existence of natural mental kinds. It is worth noting that this commitment admits of more deflationary readings. Instrumentalist or pragmatist approaches, for instance, may treat the postulation of natural kinds as serving explanatory or heuristic ends without strong realist commitments (cf. Brzović 2024; Dupré 1993). For dialectical convenience—and because the simulationist program is, as we will see, most naturally read as adopting a broadly realist orientation—we bracket these alternatives here. The central issues we identify below would arise *mutatis mutandis* within such alternative frameworks.

Second, empirical science is the best guide to uncovering the world's natural-kind structure. This view—now so entrenched in analytic philosophy as to constitute received wisdom—can be traced back at least as far as Whewell (1840) and Mill (1843), who regarded scientific taxonomies as aimed at discovering natural or real kinds. In his seminal contribution, Boyd (1991) highlighted the connection between such discovery and explanatory success: "Kinds useful for induction or explanation must always "cut the world at its joints" in this sense: successful induction and explanation always require

that we accommodate our categories to the causal structure of the world" (p. 139). Natural kinds—such as, presumably, *gold, mammal,* or *perception*—are apt for induction and explanation, and perhaps even productive theorizing more generally, because they allow us to accommodate our categories to the world's causal structure. For better or worse, it is now commonplace that the empirical sciences alone can achieve such accommodation. Unlike empirical inquiry, a priori intuition and conceptual analysis are rarely taken seriously as means of identifying natural kinds.[2]

Third, and relatedly, in investigating natural phenomena, philosophy should take its lead from the best available science. Recent work on methodological naturalism has developed this idea in several ways, which may diverge in important respects (see, e.g., Maddy 2007; Collins 2015; Ritchie 2022). One common element, important for our purposes, concerns metaphysical inquiry.[3] When investigating the nature of phenomena—e.g., mental states, events, or processes—philosophy should be guided by the categories and classification schemes of the most successful scientific theories. Positively, this requires philosophical engagement with empirically productive theories and research programs in the relevant domains. Negatively, it justifies principled suspicion of categories misaligned with those of mature scientific frameworks. In pursuing knowledge of the natural world—and of phenomena as candidates for natural kinds—there is no distinctively philosophical domain of inquiry.[4]

This does not entail that philosophy and science are always engaged in the same explanatory enterprise. Philosophical views and discussions may, for example, aim to shed light on the way things in a given domain—individuated in terms of broader theoretical or practical interests—hang together (Sellars 1962; Dennett 2015; Thomasson 2023). Philosophical views may articulate the scientific image of a phenomenon of interest, elucidate our 'manifest' (pre-scientific) image of it, or relate the two.[5] Philosophical theorizing of this kind—what we may call *Sellarsian Orientation*—is widespread in

---

[2] This does not entail that intuitions do not play *any* role in philosophical or scientific inquiry about natural kinds. The naturalist may accommodate, albeit somewhat uncomfortably, Kripke's (1980) idea that modal intuitions support *a posteriori* necessity claims about kinds.

[3] In what follows, we will frequently talk of metaphysics and ontology. We take ontology to be a branch of metaphysics that deals with questions concerning existence. Metaphysicians of the mind, as we will see, typically strive to formulate a *cognitive* ontology, identifying the kinds of existing cognitive entities and processes (see, e.g., Khalidi 2023).

[4] For balanced criticisms of this view, see, e.g., Baker (2015) and Jaksland (2023).

[5] Sellars (1962, p. 35) famously characterized the aim of philosophy as one of understanding "how things in the broadest possible sense of the term hang together in the broadest possible sense of the term", a characterization many have appealed to in attempting to accommodate such things as free will, consciousness, or indeed memory, to the scientific image of the world.

contemporary analytic philosophy and, as we will see, has been adopted in a particular form by simulationists. For the naturalist, however, the legitimacy of such theorizing does not license a distinctively philosophical method or domain of inquiry independent of, or insulated from, empirical investigation. Nor does it justify divergence from the classificatory schemes of our best scientific theories—at least not in explanatory contexts where the aim is to understand the natural world or our intuitive conceptions of it. In such contexts, the Sellarsian task presumes the authority of mature scientific theories, whether in physics, cognitive science, or comparative anthropology.

2.2. Simulationism and memory systems as natural kinds

Simulationism aims to examine a phenomenon—*episodic memory*—hypothesized to instantiate a natural kind. The examination is grounded in an empirically successful research program in the memory sciences. This program—along with several influential theories that emerged within it, particularly during the first decade of the 21st century—has yielded a novel picture of episodic memory as essentially simulative and mechanistically linked to imagination. The resulting picture challenges many traditional assumptions about memory's function and epistemic status.[6] In the opening chapter of his agenda-setting *Mental Time Travel,* Kourken Michaelian makes clear that simulationism's central goal is to render this emerging view explicit:

> Ultimately, the goal is not to develop a theory that can necessarily be adequately expressed in the form of an analysis, but rather to *formulate a useful general framework for thinking about human memory, one that draws out and makes explicit the vision of memory implicit in current psychology* and that can in turn contribute to interpreting the findings of and suggesting new lines of inquiry for the latter (2016a, p. 3, emphasis added).

Simulationism aims to synthesize the new class of theories of episodic memory, systematize their shared commitments, and bring them into contact with traditional philosophical questions (cf. Andonovski & Michaelian 2024, pp. 255–259).

It is helpful to examine the *multiple memory systems* (MMS) research program to clarify the empirical grounding of the simulationist framework. The program posits a plurality of memory

[6] In the rest of the article, we will frequently use 'memory' to refer to *episodic memory,* reserving judgment about whether episodic memory is itself a neurocognitive system. In the same spirit of neutrality, we use 'remembering' to refer to an occurrent mental process, characterized by several phenomenal and functional features (see Andonovski, forthcoming; Openshaw, forthcoming). In these terms, episodic memory partly underpins remembering.

systems, each with distinct information-processing functions, operating principles, and neuroanatomical bases (Schacter & Tulving 1994; Squire 2004). These systems enable organisms to acquire, store, and retrieve information for a range of cognitive and behavioral tasks. Memory systems, and their component processes, constitute the basic natural kinds in the domain of memory.[7] As such, they serve as suitable targets for inductive generalization and explanatory theorizing. Simulationism inherits this system-centric view: "The approach taken here departs from [philosophical] tradition by basing the classification of kinds of memory on research on distinct but interacting memory systems conducted in psychology and neuroscience" (Michaelian 2016a, p. 17). Kinds of memory, and mental kinds more generally, are identified by reference to the classification schemes adopted by empirically successful theories of memory systems and their constituent processes.

Just as simulationism accepts cognitive psychology as the appropriate foundation for philosophical theorizing about memory, so too it adopts skepticism towards categories that do not fit such classification schemes. Philosophers may, of course, categorize memory states and processes in a variety of ways, relative to broader theoretical and practical interests. However, such ecumenical pluralism does not extend to explanatory contexts where the aim is to identify phenomena as candidates for natural kinds.[8] In such contexts, classificatory flexibility gives way to more stringent ontological commitments, and the legitimacy of a category is judged by its alignment with the taxonomic structure of mature cognitive science. Accordingly, when theorizing about the cognitive systems that underlie memory—and, as we will see in a moment, other ostensibly distinct mental kinds—"we should not draw distinctions between mental states or processes where none is to be found at the neurocognitive level" (Michaelian 2016b, p. 76).

Episodic memory has held a particularly prominent position among the various memory systems identified in the multiple memory systems literature. Tulving (1983) originally characterized episodic memory as a system for recollecting previously experienced events, marked by a distinctive sense that they belong to the personal past of the remembering subject. On Tulving's early account, a key feature of the episodic memory system is the retention of information derived from first-hand experience. Yet this view of episodic memory as bearing a privileged connection to the personal past

[7] Some theorists (e.g., Michaelian 2011) have argued for the related claim that memory, constituted as it is by multiple systems with different principles of operation, is not itself a natural kind. We reserve judgment on whether this claim is entailed by the system-centric view of memory.

[8] As we will see in section 4, this is a *very* strong claim that is not universally shared even among naturalists committed to system-based or mechanistic accounts of mental kinds.

was challenged by later developments. Tulving's (1985) reflections on the case of K.C., whose amnesia—resulting from extensive brain damage, including significant bilateral hippocampal lesions—revealed a striking parallel between impairments in remembering events in one's personal past and imagining events in one's personal future. These reflections were reinforced by a growing body of behavioral and neuropsychological evidence demonstrating a close link between remembering and imagination (Atance & O'Neill 2001; Klein et al. 2002; D'Argembeau & Van der Linden 2004, 2006; Hassabis et al. 2007). Neuroimaging studies have provided further support, suggesting that remembering and imagining activate a common set of brain regions—most notably, those comprising the so-called *default mode network*, which includes the medial and lateral temporal lobes, medial prefrontal cortex, inferior parietal lobule, and posterior cingulate cortex (Okuda et al. 2003; Addis et al. 2007; Mullally & Maguire 2014).

At the core of this emerging theoretical orientation is a bold proposal: *a common neurocognitive system underlies both remembering and imagination*. This proposal is endorsed by researchers who otherwise diverge in how to characterize the function of the system in question—variously describing it as mental time travel (Tulving 2005; Suddendorf & Corballis 2007), constructive episodic simulation (Schacter & Addis 2007; Addis 2020), or scene/scenario construction (Hassabis & Maguire 2009; Cheng et al. 2016). A more serious issue concerns the interpretation of the view that remembering and imagination share a common system. Systems are typically characterized as complex neurocognitive structures with multiple, frequently interacting, component processes (Sherry & Schacter 1987; Schacter & Tulving 1994). Remembering and imagining clearly rely on "many of the same processes" (Addis & Schacter 2012, p. 1). Yet such reliance does not rule out that the common system contains specific processes that are specialized for remembering and for imagination, respectively. (For some evidence in this direction, see Addis et al. (2009), Andrews-Hanna et al. (2010), and Menon (2023.) Nevertheless, a stronger interpretation of the common system proposal has gained traction.

2.3. Simulationism and identity-of-process

On the *identity-of-process* view, one and the same neurocognitive process exhaustively underlies both remembering and imagining.[9] The process, typically characterized as *simulative* in nature, is believed

[9] We add 'exhaustively' to rule out an interpretation on which this underlying process is common to remembering and imagining but that some other specialized mnemic process (e.g., engram-retrieval) forms an additional underlying constituent of remembering. For expositional convenience, we use *'process'* to refer to what is likely a set of interacting processes. How cognitive processes are type-individuated remains an open question,

to involve the construction of event representations using information of various sorts and from various sources—not only direct experience of the target event. The outputs of the process can constitute recollection of previously experienced events but also the imagining of novel, unexperienced events pertaining to one's future or to what might have been. [10] Addis (2018) presents the identity-of-process view particularly clearly:

> From a neurocognitive perspective, it is likely that an event representation is physically instantiated in the brain in *exactly the same way irrespective of whether it is remembered or imagined*: it is a set of connections between the nodes representing perceptual content, semantic information and related schemas… [In remembering and imagining events] *the same simulation process is engaged and the same types of information are drawn upon* (p. 70, emphasis added).

One and the same simulation process wholly constitutes both remembering and imagining.[11] While remembering and imagining may differ in some ways, these differences are best explained by the selective tuning of this process to the specific demands of each task—*not* by the existence of distinct processes or mechanisms.

Michaelian has made the commitment to identity-of-process especially explicit, taking simulationism to involve the claim that "there is no difference, at the level of natural kinds, between the process that produces our representations of the events of the personal past and the process that produces our representations of the events of the personal future" (2024, p. 18). On identity-of-process, remembering *just is* the simulation of a past personal event by the relevant system, which Michaelian—seeking to remain neutral among competing account of its function—characterizes as an *episodic construction system* (ECS). No further process-level distinction between memory and imagination is

---

complicated by the terminological variability with which 'system', 'process', and 'mechanism' are deployed across the literature. Competing proposals appeal variously to algorithmic structure (Lyons 2019), mechanistic organization (Craver 2007), and task-based functional roles (Burnston 2021). Of course, the substantive content of identity-of-process depends on such criteria. A thesis of process identity presupposes a principled way of counting processes *as* the same. The dilemma developed below, however, does not require us to take a stand here: it follows from the simulationist's core commitment that, whatever scheme is adopted, the result is one process rather than two. The burden of supplying clear individuation criteria falls to the simulationist embracing horn 1 of our dilemma.

[10] Simulationist accounts, in fact, often argue that the same process underlies a number of mental activities in addition to remembering and imagination (see, e.g., Hassabis & Maguire 2009). For convenience, we focus on remembering and imagination.

[11] See also Mahr (2024): "Even though episodic simulation can produce different kinds of representations, there is no deep, categorical difference between "remembered" and "imagined" event representations in terms of how they are generated—they share the same neural implementation and cognitive mechanisms" (p. 227).

posited or entertained. The commitment to identity-of-process, indeed, underpins the canonical formulation of the simulation theory, as presented in *Mental Time Travel:*[12]

> *S* remembers an episode *e* iff:
>
> - *S* now has a representation *R* of *e*
> - *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.
>
> (Michaelian 2016a, p. 107)

An episodic construction system functions properly just in case it reliably produces accurate representations.[13] As these are employed across different tasks, the system's *aim*—what it is trying to simulate—determines the relevant criteria for success. When a properly functioning ECS aims to simulate an event from the subject's personal past, the subject is remembering—not, e.g., imagining the future or engaging in episodic counterfactual thought. Remembering is ECS simulation of an event in one's personal past.[14] Importantly, while the precise meaning of 'aim' remains unclear, on the identity-of-process view, aims are *not* used to individuate distinct ECS processes, but rather the task-sensitive modulation of the univocal simulation process.[15]

Michaelian (2016a, 2016b, 2016c) has leveraged the identity-of-process view to argue for three theses. Given naturalist commitments, *continuism* is entailed directly by identity-of-process. If natural mental kinds are individuated in terms of the neurocognitive processes that underlie them *and* the same process underlies both remembering and imagining, then the two mental activities are

---

[12] Michaelian (2016a) uses 'event' and 'episode' interchangeably. We do the same in this paper.

[13] While Michaelian (2016a) typically treats 'properly functioning' and 'reliable' as coextensive, the relationship between these two theoretical notions remains an open question. We also set aside the issue of whether Michaelian's definition is intended to provide an account of *which* episode is remembered or whether the system's proper functioning is supposed to be evaluable independently of the representation being referentially successful (see Openshaw & Michaelian (2024)).

[14] Michaelian (2024) radicalizes simulationism further by abandoning the restriction to events in one's *personal* past. For convenience, we focus on the 2016 version of simulationism here. The points we develop nevertheless apply, *mutatis mutandis,* to this more radical simulationism as well.

[15] Michaelian (2016a) characterizes his appeal to the system's aim as a "shorthand for talk of the system responding to given retrieval cues provided by either the agent or his environment" (p. 113), a characterization which is clearly provisional and, as Andonovski & Michaelian (2024) admit, may not survive translation to a scientific idiom. We bracket this issue here.

continuous. That is, they *just are* the same natural kind, namely *ECS simulation*.[16] While some differences between them may result from the modulation of the simulation process, these differences are of degree, not kind. Philosophical intuitions to the contrary are misleading and not to be trusted insofar as they fail to engage with empirical developments. Accordingly, the naturalist simulationist suspects that the threat from *discontinuism*—the view that remembering and imagining are different in kind (Perrin 2016)—is not "particularly serious, suggesting that we base our understanding of the relationship between remembering the past and imagining the future on a demonstrably productive empirical research program, rather than on ungrounded a priori intuitions" (Michaelian 2016b, p. 67).[17]

Identity-of-process also appears to be behind the simulationist's rejection of *causalism*. If remembering is identical with a process of the ECS, then a causal connection (of the causalist's favored 'appropriate' kind) to the remembered event is no longer necessary. Indeed, advocates of identity-of-process offer empirical support for the claim that the ECS can simulate a past event without using information causally derived from the subject's original experience of the remembered event. Since remembering and imagining are wholly constituted by the same process and draw upon the same types of information, the presence of a causal connection to a past event cannot be what secures the reliable representation of the past. When simulating a past event, the ECS will often use information from the original experience, but it need not do so. In many cases, the system can generate accurate representations of a past event relying solely on information from other sources—experiences of similar events, general conceptual knowledge, testimony, and so on (Michaelian 2016a; Andonovski & Michaelian 2024; Openshaw & Michaelian 2024).

Finally, simulationism holds that the reliability of the episodic construction system is necessary for successful remembering.[18] For a subject to remember successfully, they must entertain a representation generated by a properly functioning ECS. *Mnemic reliabilism* allows the simulationist to characterize memory errors in terms of ways the simulation process, and associated metacognitive

---

[16] Somewhat puzzlingly, Michaelian often characterizes ECS simulation as episodic *imagination,* speaking indeed of episodic memory as a form of imagination (e.g., 2016a, pp. 120–121). We consider this characterization unfortunate and unhelpful; thus, we try to avoid it as much as possible in this paper.

[17] De Brigard (2023), while favoring an inclusive approach akin to reflective equilibrium, expresses similar skepticism about the appeal to intuitions and the use of 'semantic ascent' in the philosophy of memory.

[18] Michaelian often seems tempted to characterize the reliability of the ECS as also a *sufficient* condition for remembering success, a characterization which may in fact adequately reflect simulationism's broader theoretical orientation. We do not examine this issue here.

processes, malfunction. In Michaelian's (2016c) taxonomy, for example, confabulation is linked to ECS unreliability, while misremembering is attributed to representational inaccuracy—a framework further developed in later work (e.g., Michaelian 2020). We can think of this component of the simulationist program as part of a broader Sellarsian Orientation: whereas theory proper articulates the emerging scientific view of episodic memory, the reliabilist account of memory error relates this scientific view to traditional issues in epistemology and philosophy of mind. The account is expected to validate some pre-theoretical intuitions about memory success—but not all. Where such intuitions conflict with the account, they should simply give way.

In summary, identity-of-process simulationism employs a naturalist methodology to advance the thesis that a single episodic construction process underlies remembering and imagining. This thesis grounds its *continuism*, *anticausalism*, and *mnemic reliabilism*, yielding a framework that aims to synthesize recent empirical findings while challenging traditional assumptions about memory's nature, function, and epistemic roles.

3. A dilemma for simulationism

This section develops the central dilemma for simulationism. In 3.1, we examine the consequences of endorsing identity-of-process, show how identity-of-process challenges the status of remembering as a natural kind, and explain how identity-of-process severely complicates simulationism's explanatory ambitions. In 3.2., we consider simulationism without identity-of-process. Without identity-of-process, simulationism becomes a discontinuist view compatible with various forms of causalism, calling into question the theory's status as a genuinely distinct option in the theoretical landscape—not least one radical and rich in philosophical implications.

3.1. Horn 1: Identity-of-process

If the same episodic construction system (ECS) process exhaustively underlies remembering and imagination (or, better, event simulation), then remembering—understood as a capacity *for* representing the personal past—is not a natural kind. An identity-of-process simulationist view is thus eliminativist about remembering. Eliminativism about remembering permits neither a theory of remembering *qua* natural kind nor a substantive account of successful remembering in terms of process reliability. Consequently, the version of anticausalism that the view supports is of limited theoretical interest: in denying that remembering is a natural kind, identity-of-process simulationism dissolves rather than resolves the debate over the causal profile of remembering.

To clarify this line of argument, it is helpful to start with the simulationist characterization of the episodic construction system. As noted above, simulationism is officially neutral about the precise function of the system—neutral both with respect to its core computational operations and their neural implementation (Michaelian 2016a, pp. 105–106).[19] The rationale for such neutrality is straightforward. Simulationism aims to articulate the emerging scientific understanding of episodic memory. As such, it is designed to be compatible with the competing scientific theories of the underlying system, including mental time travel, constructive episodic simulation, and scene construction. At the same time:

> [The account] is meant to be sufficiently determinate to capture the common core of these approaches, the claim that *episodic remembering is not the product of a dedicated episodic memory system, but rather one product of a system responsible for episodic imagination more generally* (p. 106, emphasis added).

On this interpretation, the common core of the various competing scientific theories consists of two closely related components. First, remembering is a product of a broader episodic construction system that supports multiple—intuitively distinct—mental activities, including future-oriented imagining. Second, whatever the precise function of the episodic construction system is, it is *not* remembering, but something broader, such as event simulation. Because the function of the ECS is not remembering, the identity-of-process simulationist account of remembering appeals instead to the so-called *aims* of the ECS. Yet on identity-of-process simulationism, the aim of representing an event from the subject's past does not correspond to a dedicated process of the ECS. Rather, insofar as remembering is taken to involve such representation, there is no ECS process *for* remembering. These observations form the basis of the central argument of this section—an argument that makes explicit the eliminativist implications of the identity-of-process view:

(1) (*No dedicated process*) The ECS does not contain a dedicated process for representing events from the subject's personal past.

(2) *(Remembering$_{DEF}$)* Remembering constitutively involves representation of events from the subject's personal past.

(3) Therefore, the ECS does not contain a dedicated remembering process. (From 1, 2)

(4) *(Simulationist naturalism)* Cognitive natural kinds are individuated by the underlying neurocognitive systems and their component processes.

[19] De Brigard (2014) is *not* neutral, providing a broad characterization of the system as one for 'episodic hypothetical thinking', a characterization he significantly qualifies in later work (2023, 2025).

(5) Therefore, remembering is not a cognitive natural kind. (From 3, 4).

Premise (1) is a direct consequence of identity-of-process and has been defended explicitly by simulationists. Premise (2) articulates a constitutive feature of remembering. It does not aim to offer a full definition, but only a minimal constraint on the meaning of *remembering*—as opposed to merely imagining or simulating. The premise is intuitively plausible and nearly universally accepted in the literature. Moreover, it plays a central role in the simulationist account: it motivates their appeal to a distinct 'aim' of the ECS and is implicit in their commitment to mnemic reliability as well their account of successful remembering. We acknowledge the possibility that premise (2) could be challenged by empirical developments leading to revisions in both theoretical commitments and the meaning of *remembering*.[20] Nonetheless, our argument does not depend on fixing the term's usage. Our aim is not to establish that no natural kind could be labeled *remembering*—e.g., by a successful revisionist theory. Rather, it is to diagnose the conceptual consequences of identity-of-process for the status of remembering as currently understood.

Premise (4) articulates the naturalist commitment to individuating mental kinds in terms of the underlying systems and processes identified by successful scientific theories. Together with (3), it supports the conclusion of the argument. If the ECS has no dedicated remembering process, and natural kinds are individuated by such underlying processes, then remembering is not a natural kind. This conclusion, while surprising, has been articulated recently by Andonovski et al. (2024), who writes:

> The multiple systems approach had treated memory systems and their constituent processes as constituting the basic memory kinds, which afford productive and inductively fruitful theorizing. But if, in contrast, the identity-of-process hypothesis held by common system theorists is correct, then episodic memory is *not* a real kind; indeed, it is not even a sub-kind of mental time travel or constructive episodic simulation. As a result, there are no causal generalizations about episodic memory that carry inductive potential (p. 4).

The goal of Andonovski et al. is to draw attention to a significant shift in theorizing about the system hypothesized to underlie remembering. Our goal is to evaluate the consequences of this shift for philosophical simulationism and its associated theses, particularly in light of its naturalist commitments.

[20] Such a revision may indeed underlie Michaelian's (2024) radicalization of the simulationist theory.

Simulationism aims to systematize and make explicit the scientific image of episodic memory, under the assumption that episodic memory is a natural kind, apt for induction and explanation. If the identity-of-process view is correct, however, the assumption that episodic memory is a natural kind is unwarranted. The emerging scientific image depicts a cognitive system for event simulation, with no dedicated processes for remembering or imagining. Hence, a simulationist account that treats episodic memory as a natural kind—a kind that merits explanation in its own right—fails to capture the common core of systems theories and does not properly articulate the 'vision of memory implicit in current psychology'.

Identity-of-process simulationism is a recognizably *continuist* view. A single ECS process, hypothesized to instantiate a natural kind, underlies both remembering and imagination. In fact, if we accommodate the world's causal structure, both remembering and imagination should be dispensed with—at least in explanatory contexts aiming at revealing this structure. Since "we should not draw distinctions between mental states or processes where none is to be found at the neurocognitive level" (Michaelian 2016b, p. 76), we should not distinguish between remembering and imagining. We should instead speak of 'event simulation'. This conclusion, of course, does not preclude us from drawing the distinction for practical, normative, or epistemic purposes; it simply denies that these categories map onto distinct natural kinds (cf. Michaelian 2024, pp. 17–18).[21]

Yet, continuism comes at a cost. Simulationism aims to account for successful remembering in terms of the reliability of the ECS system. But if there is no ECS process dedicated to representing events from the subject's personal past, there is no process the reliability of which can be assessed to characterize successful remembering. If there is no remembering, there is no successful remembering. Likewise for memory failure: the ECS simulation process may malfunction in ways amenable to scientific explanation, but since ECS is not specialized for remembering, such failures are not errors of remembering—neither in an intuitive nor a theoretically satisfying sense. *Mnemic reliabilism* looks unsupportable.

The simulationist may reply that, relative to the *aim* of the ECS process, such malfunctions may constitute errors of remembering. But this route is not available under the assumption of function neutrality. If we do not know what the function of the ECS is, then we cannot identify instances in which the ECS performs the function (im)properly (see Robins (2022); Schwartz (2020)). Talk of aims

[21] For similar concerns about the collapse of simulationism into eliminativism, see Hoerl (2022).

merely smuggles an assumption that the system is trying to remember in cases independently identified as remembering. In addition, some common system theories assign the ECS an explicitly *non-mnemic* function—such as simulating possible future events (De Brigard 2014; Addis 2018). On these theories, apparent remembering errors often reflect the successful functioning of a system "that is actually doing something else" (De Brigard 2014, p. 167). Crucially, these may occur even when subjects *take themselves* to be remembering.

On the identity-of-process view, there is no reason whatsoever to expect that a properly functioning ECS will reliably produce accurate representations *of the subject's personal past*. This conclusion frustrates Michaelian's (2016c, 2020) attempts to develop a taxonomy of memory errors in terms of ECS failures. It also provides a reason to be extremely suspicious of the simulationist variant of mnemic reliabilism: if the ECS is not for remembering, its reliability is not sufficient for accurate or successful remembering.

A similar dialectic plays out in the debate about causation. Simulationist *anticausalism* is grounded in a claim about the function of the ECS: the system can simulate a past event without accessing information derived from the subject's experience of the target event. This may well be true of the ECS, but it is misleading to frame it as a substantive insight about remembering *per se*. On identity-of-process, no class of states instantiates a natural kind that corresponds to the pre-theoretical category of remembering. Hence, there are no informative generalizations to be made about the causal profiles of remembering states. Simulationists appeal to details about the function of the ECS to argue that a causal connection to a past event is not *necessary* for remembering. But establishing the lack of necessity requires showing that, across all instances of remembering, causation is absent in at least some cases. And showing that presupposes states of remembering that can be identified independently of their causal profiles—something the identity-of-process view denies explicitly. If remembering is not a natural kind, then we lack the theoretical basis for drawing principled distinctions between genuine and merely apparent instances of remembering.

At this point, simulationists may insist that anticausalism remains insightful. We regularly identify instances of remembering for practical or epistemic purposes even if we do not consider them tokens of a distinct natural kind. It is informative and surprising that the system responsible for generating these states does not require or privilege information from the subject's past experience of the target event. We agree. Indeed, our analysis of the identity-of-process view underscores how radical and surprising it is—how far it departs from our pre-theoretical expectations. However, to properly

appreciate the force of the identity-of-process view, it should not be understood as offering a philosophical theory of remembering, alternative to and on a par with causal or epistemic theories (cf. Michaelian & Sutton 2017). Rather, it should be seen as challenging the very assumption that remembering is a mental kind that admits of theorizing in the first place. Simulationism, properly understood, dissolves rather than resolves the causalism-anticausalism debate.

Horn 1 simulationism—i.e., identity-of-process simulationism—forces us to confront the radical nature of systems-centric theorising. It entails eliminativism about remembering (and imagining) and makes debates over causation and memory success moot. However, this is not the only way forward. Another option, to which we now turn, abandons the commitment to identity-of-process.

3.2. Horn 2: Difference-of-process

If, by contrast with the identity-of-process thesis, remembering is a product of a dedicated ECS process, then remembering *is* a natural kind. A simulationist view that posits such a process would be in a position to provide a substantive account of remembering in terms of process reliability. Such a version of simulationism would be *discontinuist*: it would reject the core idea of the common systems theories described in §2.2.

As we have seen, Michaelian and Addis openly endorse identity-of-process. Nevertheless, elements of a weaker interpretation of the common system view can be found in their work. Michaelian (2016a) tells us that remembering *is* "a constructive process intimately related to future-oriented mental time travel and to a broader range of forms of episodic imagination" (p. 57). As such, it is "not different in kind from *other* episodic constructive processes" (p. 103, emphasis added). Imagining a novel event, unlike remembering, may nevertheless require *recombination* of episodic details during retrieval—a process Addis & Schacter (2012) take to be distinctive of imagination.[22] De Brigard, in contrast, has been reluctant to endorse a strong identity-of-process view, speaking only of "common neural mechanisms" underlying memory and imagination (e.g., 2023, p. 44) and characterizing remembering as "just one operation" of a larger cognitive system for episodic hypothetical thought (2014, p. 158). In light of this—and given the apparently radical nature of the identity-of-process view—

[22] Addis (2018) abandons this view, arguing that the relevant difference between memories and imaginings reflects a difference in 'associative history' and thus does not require positing a recombination process proprietary to imagination. In a recent article, Mahr & Schacter (2024) characterize episodic recombination as a natural kind of compositional computation that underlies *all* forms of ECS simulation.

it would be theoretically fruitful to articulate a weaker version of simulationism, to assess how it fares, both with respect to its core commitments and with regard to the associated theses.

This alternative, weaker version of simulationism takes the appeal to a distinct *aim* of the ECS system seriously by positing a dedicated ECS process for remembering. *Ex hypothesi*, the function of this process is to accurately represent episodes from the subject's personal past. This commitment invites a revision of the canonical formulation of the simulation theory. On Horn 2 simulationism:

> *S* remembers an episode *e* iff:
>
> - *S* now has a representation *R* of *e*
> - *R* is produced by a properly functioning process $P_{remembering}$ of the episodic construction system.

Remembering is underlaid by a process $P_{remembering}$, whose function is distinctively mnemic. The process is nevertheless a constituent operation of the ECS and interacts frequently with other processes of the system—e.g., processes specialized for the imagining future or counterfactual events. Remembering, on this view, is a *kind* of simulation—one with a distinctively mnemic aim.

This version of simulationism restores the appeal of *mnemic reliabilism*. A dedicated ECS process for representing the personal past provides a basis for theorizing about, and distinguishing between, successful and unsuccessful remembering. The reliability of $P_{remembering}$ can be used to articulate conditions for successful remembering as well as for categorizing different forms of memory error—such as distortions, omissions, or confabulations—as kinds of process malfunctions. Since $P_{remembering}$ has a mnemic function, its reliability can be assessed in terms of how accurately its products represent the subject's personal past. That said, the reliabilist account is grounded not in the accuracy of particular memory representations, but in the proper functioning of the process that produces them. For the naturalist, appeals to proper function are underscored by the possibility of accidentally accurate representations: even a malfunctioning system—like a broken watch—can occasionally produce correct outputs, but such outputs do not constitute remembering in a *theoretically relevant sense* (cf. Michaelian 2016c; Robins 2020). They are underscored as well by the possibility of non-accidentally inaccurate representations generated by a process that is functioning properly: landing in Berlin on a flight from New York your watch may get the time wrong, but not because it is broken.

The simulationist who rejects identity-of-process is in a position to develop a taxonomy of memory errors in terms of ECS malfunctions and predictable misfires. Yet, they must proceed cautiously. The correspondence between kinds of system malfunction and misfire on the one hand,

and pretheoretically identified memory errors on the other, may be weak or moderate. Michaelian (2016c), for example, characterizes *misremembering*—an apparent kind of memory error in which subjects mischaracterize an event in some respect, despite retaining much information about it—as a byproduct of a properly functioning memory system. On this view, misremembering may be considered an error, but only in a theoretically attenuated sense.[23] Yet it is neither intuitively obvious nor empirically well-established that in all cases of misremembering—e.g., in the so-called DRM paradigm (Roediger & McDermott 1995)—the relevant ECS process is functioning properly. Accordingly, future work on the taxonomy of memory errors should proceed in close coordination with empirical research into the ECS, its component processes, their proper functions, and the background conditions under which they function.

Horn 2 simulationism, however, is essentially a *discontinuist* view: it posits dedicated processes for each of remembering and imagination as components of the ECS and thus marks off *sub-kinds* of episodic simulation. For the purposes of naturalist theorizing, sub-kinds *are* real kinds: they carve the episodic construction system at a joint and support causal generalizations with inductive potential. Indeed, this point has been acknowledged by participants in the (dis)continuism debate, where the existence of functionally-discrete components within the ECS has been invoked to justify treating remembering and imagining as distinct natural kinds. Discontinuists—such as Perrin (2016), Cheng et al. (2016), and Werning (2020)—recognize that remembering involves episodic simulation by a proprietary neurocognitive system. Yet they challenge identity-of-process and posit a functionally specialized process for remembering as a target of independent theoretical and empirical inquiry.[24]

Perhaps unexpectedly, discontinuist Horn 2 simulationism has the resources to ground a substantive *anticausalist* thesis. If there is an ECS process $P_{remembering}$, then its outputs constitute a natural kind that corresponds to the pretheoretical category of remembering and supports empirically tractable generalizations—including those concerning their causal profiles. On this way of framing the dialectic, discontinuism is not merely compatible with anticausalism, it is necessary for articulating anticausalism as a serious theoretical option. That said, discontinuist simulationism is also *compatible* with different forms of causalism. It is an empirical question whether states produced by $P_{remembering}$ have causal profiles that require a causal connection to the remembered event. Presumably, this

[23] After all, it may be that in cases of misremembering, "the system is doing what it is supposed to do" (De Brigard 2014, p. 172).

[24] Indeed, theorists like Werning (2020) typically treat 'simulation' as synonymous with 'scenario construction', which they take to be partly constitutive of episodic memory.

empirical question is open. Indeed, a growing number of theorists who share the simulationist's naturalist commitments advance causalist views that aim to accommodate the evidence for memory (re)construction (e.g., Werning 2020; Najenson 2021; Andonovski 2024; Cheng 2024; Schwartz 2025). To the extent that such views incorporate the idea that remembering constitutively involves simulation, they can be tentatively characterized as both causalist *and* simulationist (in the Horn 2 sense).

One might expect the simulationist to insist that we lack evidence that a causal connection of the relevant sort is necessary, and that the onus is on causalists to justify such a requirement. Yet what such evidence for necessity would be is a difficult question that has yet to be addressed systematically by memory theorists (but see Werning 2020; Andonovski 2021, forthcoming; Schwartz 2025). For now, it is worth noting that the simulationist's signature argument for anticausalism typically *invokes identity-of-process:*

> Since imagining a future event trivially does not presuppose the existence of a causal connection between the subject's thought of the event and his experience of the event, this, in turn, suggests that *remembering a past event likewise should not be taken to presuppose the existence of a causal connection* between the subject's thought of the event and his experience of the event (Michaelian & Sant'Anna 2021; p. S320, emphasis added).

This argument simply does not get off the ground if there *are* distinct ECS processes for remembering and imagining. If we assume that $P_{remembering}$ and $P_{imagining}$ differ in function, we should forbid direct inferences from the causal profile of representations produced by one to the causal profile of representations produced by the other. Accordingly, if identity-of-process fails, defending anticausalism requires more direct evidence about the causal profile of remembering states and the functional characteristics of the process responsible for producing them.[25]

In sum, Horn 2 simulationism abandons identity-of-process, positing a dedicated process for remembering. In doing so, it becomes a more modest, less revisionary theory. But it also becomes untethered to its stated philosophical ambitions, such as anticausalism. Horn 2 simulationism gambles with the core ideas that make simulationism distinctive; whether it constitutes a distinct option in the theoretical landscape—not least one radical and rich in its philosophical implications—is an open empirical question.[26]

---

[25] For a simulationist anti-necessity argument that appeals to the empirical possibility of generating an entirely new memory representation at retrieval, see, e.g., Michaelian & Sant'Anna (2021, pp. 7483–7484).

[26] For relevant discussion that has emerged since this paper was first drafted, see Álvarez (2026).

3.3. Summary of the dilemma: Eliminate or refine

There are two apparent options for simulationists committed to individuating memory and imagination in terms of the neurocognitive systems and component processes that underlie them. The first posits a single simulation process that exhaustively underlies both remembering and imagining. This identity-of-process view eliminates remembering as a natural kind and thus, by the simulationist's lights, as a target for productive theorizing.[27] The second posits a dedicated simulation-like process, or set of processes, *for* remembering. This discontinuist view tempers the revisionary ambitions of simulationism and is compatible with both causalist and anticausalist theories of remembering. The two options should be familiar from the broader debates about cognitive ontology, with theorists often facing the choice between eliminating (folk)psychological kinds and grounding them in more specific, fine-grained processes (Poldrack 2010; Burnston 2021; McCaffrey & Wright 2022). Yet, as these debates illustrate, there are other options on the table. In section 4, we briefly examine two principal strategies for escaping the dilemma.

4. Out of the dilemma?

In 4.1, we consider a version of simulationism that relaxes the commitment to individuating mental kinds in terms of underlying neurocognitive systems and explore how this shift affects the explanatory ambitions of simulationism. In 4.2, we examine a broader reconstrual of simulationism in a naturalist framework that does not rely on any mental kind hypotheses and characterize it as a form of Sellarsian Orientation.

4.1. Beyond system-centrism

The methodological basis of simulationism includes a commitment to the multiple memory systems (MMS) hypothesis. For MMS theorists, and consequently for simulationists, memory systems are the

[27] As we flagged in fn. 8, the strong tie between natural kindhood and productive theorizing is controversial. Functional, folk-psychological, and task-based modes of theorizing have all been defended without requiring that their target phenomena instantiate natural kinds (see, e.g., Burnston 2021; McCaffrey & Wright 2022). Our claim is not that this tie holds. Indeed, we ourselves are suspicious of such a strong claim. Rather, our point here is that the simulationist, as reconstructed in §2, is explicitly committed to such a claim. See, for instance, Michaelian's (2016b, p. 76) injunction that "we should not draw distinctions between mental states or processes where none is to be found at the neurocognitive level". The eliminativist conclusion thus follows from the simulationist's own methodological commitments, not from any thesis endorsed by the authors. In §4, we consider what becomes available for simulationists when some of those commitments are relaxed.

basic natural kinds in the domain of memory. As such, they are the proper focus of explanatory theorizing, posited to explain the "introspectively apprehensible and objectively identifiable consequences of learning and memory" (Schacter & Tulving 1994, p. 3). Moving beyond system-centrism would significantly alter the simulationist approach. Still, it may be possible to remain within the bounds of philosophical naturalism and preserve simulationism's central insight—that memory and imagination are intimately related—without embracing the claim that mental natural kinds must be individuated in strict accordance with the systems posited by prevailing scientific theories.

One option abandons the assumption that systems theories offer the only legitimate framework for individuating mental kinds. Doing so undermines confidence not only in the standard model of memory systems (cf. De Brigard 2023), but also in systems theorizing more generally. This route has been pursued by theorists with markedly different explanatory goals and approaches and has gained popularity recently. Anderson (2014) and Pessoa (2022), for example, treat systems theorizing as a phrenological relic, and advocate for a conception of the brain as a complex, entangled organ whose functioning is supported by intricate brain-body-environment dynamics operating across many different scales. The brain is complex and brain researchers have diverse aims. Such complexity and diversity may call for taxonomical pluralism in neuroscience as well as cognitive psychology (e.g., Sullivan 2017; McCaffrey & Wright 2022).

In a recent article, Burnston (2021) argues that psychological faculties should be treated not as explanatory but *heuristic* posits that provide traction on an extremely complex space of behavioral capacities. Burnston asks us to investigate how "information processing in the brain explains behavior directly, and not in virtue of instantiating some particular mental function" (p. 272). Such a program targets differences in structure across tasks to shed light on how neural function supports complex behavioral abilities—abilities such as remembering or imagining (see Figdor 2011). Adopting this perspective would be costly to the system-centric simulationist. Yet similar ideas have surfaced recently in the broader simulationist literature. De Brigard (2025) suggests that the default mode network constitutes not a unified system whose function can be characterized in psychological terms, but a more basic metabolic processes (see also Raichle et al. 2001). Likewise, Mahr (2020, 2024) recognizes that common or overlapping mechanisms are frequently deployed in tasks across seemingly heterogeneous psychological domains. While Mahr is happy to talk of an 'episodic simulation system', his broader aims may be accommodated within a program that dispenses with the system-centric assumptions of the MMS approach.

Simulationists who opt for this route may be reasonably confident that they have strayed only minimally, if at all, from naturalistic strictures. Indeed, simulationists who motivate their rejection system-centrism by appeal to evidence of the widespread multifunctionality of brain areas may insist that they *are* anchoring their cognitive ontology on the best available science. That such ontology has no room for a unified episodic construction system should not hinder investigation into the similarities and differences between remembering and imagining or the (common) neural mechanisms that support them. If the aim of simulationism is to make explicit the vision of memory arising from the new science, then it may be better served by recognizing the complexity and the incompleteness of that vision. Chief among these is undoubtedly the close processing connection between memory and imagination, a connection that can be recognized without a commitment to (dis)continuism or speculative claims about the causal profile of remembering states. A simulationism revised along these lines is unlikely to appeal to theorists who seek a definitive account of episodic memory or a resolution to the (anti)causalism debates. Yet the willingness to live with incompleteness and abandon the expectation of neat resolutions to ontological disputes is arguably the naturalistic attitude most worth cultivating.

4.2. Simulationism and Sellarsian Orientation

Another strategy that may be pursued independently of—or in tandem with—rejecting system-centrism, repositions simulationism within the broader theoretical landscape of naturalism. On this strategy, the scientific vision of memory is not a set of hypotheses about episodic memory or simulation as natural neurocognitive kinds, but a *general* framework that can be brought into critical relation with the manifest image of memory found in commonsense conceptions, traditional philosophical theories, and practical, ethical, or legal attitudes toward remembering.[28]

This strategy, which sees simulationism as pursuing a kind of Sellarsian Orientation, has already shaped its seemingly negative project—i.e., its criticism of causal or epistemic theories of memory. The central target of this criticism is the familiar view of memory as a faculty that affords epistemically privileged access to the personal past by the preservation and transmission of information acquired first-hand. Simulationism challenges this picture in two closely connected ways. First, it underscores the intimate relation between memory and imagination, exemplified, at a

[28] On this view, explanatory theorizing in the *sciences* can still be conceived as grounded in the pursuit of natural kinds. Indeed, the view may be particularly appealing to those seeking to carve out distinct roles for science and philosophy (cf. Sellars 1962; Maddy 2007).

minimum, by the existence of shared simulation processes. This intimate relation suggests that the epistemic profiles of remembering and imagining are tightly linked: if remembering constitutively involves constructive simulative retrieval—a point accepted by continuists and discontinuists alike—then we should expect remembering to display features characteristic of imaginative success and failure, such as coherence and plausibility on the one hand, and poorly constrained confabulation ('making stuff up') on the other. Second, and relatedly, it underscores the empirical evidence that reveals transmissionist conceptions of the sort embodied in classic causal and epistemic theories as overly simplistic, empirically suspect, and inadequate to the task of explaining the reliability of memory.[29] Even if information acquired first-hand were preserved relatively unaltered until recall, the very existence of a constructive simulative retrieval process shows that such information can account for the nature and epistemic status of remembering at best only partially.

The version of naturalism that simulationists endorse exemplifies a strong commitment to the priority of the scientific image—particularly, but not exclusively, in ontological contexts. Seen in this light, simulationism offers a general characterization of the scientific image of memory as essentially simulative and imagination-like in several epistemically significant respects. Crucially, this characterization can be developed without endorsing a robust form of anticausalism or any strong claim about constructive simulation as a neural kind without thereby diminishing its significance. We suspect that this version of simulationism will not appeal to those who seek a *theory* of memory with the formal or deductive structure characteristic of traditional causal or epistemic accounts. Yet, if simulationism is to develop as a useful general framework that "captures what is important about *human* remembering [...] as it unfolds in the real world" (Michaelian 2016a, p. 4), the Sellarsian Orientation may best realize that aim.

5. Conclusion

Simulationism reinvigorated the philosophical study of memory. As the leading 'post-causal' theory, simulationism has challenged traditional conceptions of memory as a source of privileged, unmediated access to the personal past. Yet, we have argued, simulationism exhibits a set of methodological and theoretical tensions that complicate its explanatory ambitions. At its core lies a dilemma: if it is serious

[29] *Transmissionist* theories require the storage and transmission of informational content from encoding to retrieval. Contemporary causal theories are non-transmissionist. For discussion, see Michaelian & Robins (2018) and Openshaw (2023).

about embracing the identity-of-process view, simulationism eliminates remembering as a natural kind, and thus as an object of scientific and philosophical theorizing. If, on the other hand, it abandons identity-of-process, it becomes a softer, discontinuist view, one compatible with causalism and anticausalism alike.

We do not call for a wholesale rejection of the central insights or significance of simulationism. On the contrary, we take its challenge to traditional views of memory seriously. We believe that a careful, methodologically disciplined examination of simulationism is necessary to advance philosophical reflection on memory and imagination. We believe that the future of the metaphysics of memory lies not in the continuation of narrowly-framed debates about memory causation, pursued in isolation from scientific inquiry. Rather, it lies in a more sustained engagement with emerging evidence, accompanied by an explicit appreciation of the methodological and theoretical commitments required to accommodate such evidence.[†]

[†] *Acknowledgements:* For discussions we would like to thank Juan Álvarez, Donna Addis, José Carlos Camillo, Carl Craver, Felipe De Brigard, Andrew Forcehimes, Kristina Liefke, Nic Koziolek, Denis Perrin, Louis Renoult, Markus Werning, and especially Kirk Michaelian. We are grateful to those who attended presentations of this material in Grenoble at Simulationism 2023, the 2024 meeting of the European Society for Philosophy and Psychology, and the 2nd Shanghai-Grenoble Philosophy of Memory workshop in Shanghai.

Funding: Andonovski received funding from the European Union's Horizon research and innovation program under the Marie Sklodowska-Curie grant agreement 101062754. Copenhaver received funding from the Center for the Humanities at Washington University in St. Louis and from the Centre for Philosophy of Memory at the Institut de Philosophie de Grenoble at the Université Grenoble Alpes. Openshaw's work on this project was supported by a Nanyang Technological University Start-Up Grant.

Full authorship statement: This paper was developed in conversation between the authors beginning in 2023. The first draft was prepared by Copenhaver and delivered at Simulationism 2023 in Grenoble. A second draft was prepared by Openshaw with comments from Andonovski. A second presentation was drafted by Copenhaver and delivered at the 2024 meeting of the European Society for Philosophy and Psychology. Copenhaver prepared a third draft, which was reviewed and received comments by Andonovski and then was revised substantially by Openshaw. Finally, a substantially revised fourth draft was prepared by Andonovski and presented at the 2nd Shanghai-Grenoble Philosophy of Memory workshop, with light revisions to the final manuscript by Copenhaver and Openshaw. Openshaw handled the revisions requested by the journal. The authors contributed equally.

Bibliography

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.

Addis, D. R. (2020). Mental time travel? A neurocognitive model of event simulation. *Review of Philosophy and Psychology,* 11, 233–259.

Addis, D. R., Pan, L., Vu, M. A., Laiser, N., & Schacter, D. L. (2009). Constructive episodic simulation of the future and the past: Distinct subsystems of a core brain network mediate imagining and remembering. *Neuropsychologia*, *47*(11), 2222–2238.

Addis, D. R., & Schacter, D. L. (2012). The hippocampus and imagining the future: where do we stand?. *Frontiers in human neuroscience*, *5*, 173.

Álvarez, J. F. (2026). Memory and imagination: Toward discontinuist simulationism. *Philosophical Psychology*, 1–31. DOI: 10.1080/09515089.2026.2619128

Anderson, M. L. (2014). *After phrenology* (Vol. 547). Cambridge, MA: MIT Press.

Andonovski, N. (2021). Causation in memory: Necessity, reliability and probability. *Acta Scientiarum*. Human and Social Sciences, 43(3), e61493.

Andonovski, N. (2024). Engrams as mental files. *Synthese,* 204(6), 154.

Andonovski, N. (forthcoming). Memory and Causation. In A. Sant'Anna & C. Craver (Eds.) *The Oxford Handbook of Philosophy of Memory*. Oxford University Press.

Andonovski, N., & Michaelian, K. (2024). Naturalism and simulationism in the philosophy of memory. In *Naturalism and its challenges* (pp. 252–273). Routledge.

Andonovski, N., Sutton, J., & McCarroll, C. J. (2024). Eliminating episodic memory?. *Philosophical Transactions B*, *379*(1913), 20230413.

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., and Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron* 65, 550–562.

Atance, C. M., & O'Neill, D. K. (2001). Episodic future thinking. *Trends in cognitive sciences*, *5*(12), 533–539.

Baker, L. R. (2015). Ontology Down-to-Earth. *The Monist*, *98*(2), 145.

Bernecker, S. (2017). A causal theory of mnemonic confabulation. *Frontiers in Psychology*, *8*, 1207.

Bird, A. & Tobin, E. (2024). Natural kinds. In *The Stanford Encyclopedia of Philosophy* (Spring 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.). https://plato.stanford.edu/archives/spr2024/entries/natural-kinds/.

Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 61*(1/2), 127–148.

Brzović, Z. (2024). The many faces of realism about natural kinds. *Journal for General Philosophy of Science, 55*(3), 289–307.

Burnston, D. C. (2021). Cognitive ontologies, task ontologies, and explanation in cognitive neuroscience. In J. Bickle, C. F. Craver, and A.-S. Barwich (eds.) *The Tools of Neuroscience Experiment: Philosophical and Scientific Perspectives* (pp. 259–283). Routledge.

Camillo, J. C. (2025). Interpreting continuism as a mechanistic thesis. *Synthese, 205*(4), 1–24.

Cheng, S. (2024). Distinct mechanisms and functions of episodic memory. *Philosophical Transactions B,* 379(1913), 20230411

Cheng, S., Werning, M., & Suddendorf, T. (2016). Dissociating memory traces and scenario construction in mental time travel. *Neuroscience & Biobehavioral Reviews,* 60, 82–89.

Clark, K. J. (Ed.). (2015). *The Blackwell companion to naturalism*. John Wiley & Sons.

Collins, J. (2015). Naturalism without metaphysics. In Fischer, E. and Collins, J. (eds.). *Experimental philosophy, rationalism, and naturalism* (85–109). London: Routledge.

Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford University Press.

D'Argembeau, A., & Van der Linden, M. (2004). Phenomenal characteristics associated with projecting oneself back into the past and forward into the future: Influence of valence and temporal distance. *Consciousness and cognition*, 13(4), 844–858.

D'Argembeau, A., & Van der Linden, M. (2006). Individual differences in the phenomenology of mental time travel: The effect of vivid visual imagery and emotion regulation strategies. *Consciousness and cognition*, 15(2), 342–350.

De Brigard, F. (2014). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese, 191*(2), 155–185.

De Brigard, F. (2023). *Memory and remembering*. Cambridge University Press.

De Brigard, Felipe. (2025). Simulationism and Memory Traces. In Lynn Nadel & Sara Aronowitz (eds.), *Space, Time, and Memory* (pp. 194–219). Oxford University Press

Dennett, D. C. (2015). *Elbow room, new edition: The varieties of free will worth wanting*. MIT Press.

Dupré, J. (1993). *The disorder of things: metaphysical foundations of the disunity of science*. Harvard University Press.

Figdor, C. (2011). Semantics and metaphysics in informatics: Toward an ontology of tasks. *Topics in Cognitive Science, 3*(2), 222–226.

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences,* 104(5), 1726–1731.

Hassabis, D. & Maguire, E.A. (2009). The construction system of the brain. *Philosophical Transactions of the Royal Society B* 364: 1263–1271

Hoerl, C. (2022). A knowledge-first approach to episodic memory. *Synthese* 200, 376. DOI: 10.1007/s11229-022-03702-1

Jaksland, R. (2023). Naturalized metaphysics or displacing metaphysicians to save metaphysics. *Synthese 201*(6), 199.

Kemp, G. N., Khani, A. H., Rezaee, H. S., & Amiriara, H. (Eds.). (2025). *Naturalism and Its Challenges*. Routledge.

Khalidi, M. A. (2018). Natural kinds as nodes in causal networks. *Synthese, 195*(4), 1379–1396.

Khalidi, M. A. (2023). *Cognitive ontology*. Cambridge University Press.

Klein, S. B., Loftus, J., & Kihlstrom, J. F. (2002). Memory and temporal experience: The effects of episodic memory loss on an amnesic patient's ability to remember the past and imagine the future. *Social Cognition*, 20(5), 353–379.

Kripke, S. A. (1980). *Naming and necessity* (Vol. 217). Oxford: Blackwell.

Lai, C. (2024). Remembering requires no reliability. *Philosophical Studies, 181*(1), 43–63.

Langland-Hassan, P. (2022). Propping up the causal theory. *Synthese, 200*(2), 95.

Liefke, K., & Werning, M. (2024). Diachronicity matters! How semantics supports discontinuism about remembering and imagining. *Topoi, 43*(4), 1137–1159.

Lyons, J. C. (2019). Algorithm and parameters: Solving the generality problem for reliabilism. *Philosophical Review, 128*(4), 463–509.

Maddy, P. (2007). *Second philosophy: A naturalistic method*. Oxford: Oxford University Press.

Mahr, J. B. (2020). The dimensions of episodic simulation. *Cognition,* 196, 104085.

Mahr, J. B. (2024). How to become a memory: the individual and collective aspects of mnemicity. *Topics in Cognitive Science, 16*(2), 225–240.

Mahr, J. B., & Schacter, D. L. (2024). Episodic recombination and the role of time in mental travel. *Philosophical Transactions B, 379*(1913), 20230409.

McCaffrey, J., & Wright, J. (2022). Neuroscience and Cognitive Ontology: A case for pluralism. In F. De Brigard & W. Sinnott-Armstrong (Eds.), *Neuroscience and Philosophy*. MIT Press.

Menon, V. (2023). 20 years of the default mode network: A review and synthesis. *Neuron, 111*(16), 2469–2487.

Michaelian, K. (2011). Is memory a natural kind?. *Memory Studies, 4*(2), 170–189.

Michaelian, K. (2015). Opening the doors of memory: is declarative memory a natural kind?. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(6), 475–482.

Michaelian, K. (2016a). *Mental Time Travel: Episodic Memory and Our Knowledge of the Personal Past*. MIT Press.

Michaelian, K. (2016b) Against discontinuism: Mental time travel and our knowledge of past and future events. In *Seeing the Future: Theoretical Perspectives on Future-Oriented Mental Time Travel*. (Eds. K. Michaelian, S. B. Klein, and K. K. Szpunar.) New York: Oxford University Press.

Michaelian, K. (2016c). Confabulating, Misremembering, Relearning: The Simulation Theory of Memory and Unsuccessful Remembering. *Frontiers in Psychology, 7*.

Michaelian, K. (2020). Confabulating as Unreliable Imagining: In Defence of the Simulationist Account of Unsuccessful Remembering. *Topoi* 39, 133–148.

Michaelian, K. (2024). Radicalizing simulationism: Remembering as imagining the (nonpersonal) past. *Philosophical Psychology, 37*(5), 1170–1196.

Michaelian, K., & Robins, S. K. (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. 13–32). London: Routledge.

Michaelian, K., & Sant'Anna, A. (2021). Memory without content? Radical enactivism and (post) causal theories of memory. *Synthese, 198*(Suppl 1), 307–335.

Michaelian, K. & Sutton, J. (2017). Memory. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Summer 2017). https://plato.stanford.edu/archives/sum2017/entries/memory/

Mill, J. S. (1843/1882). *A system of logic (8th edition)*. Harper & Brothers.

Moran, A. (2022). Memory disjunctivism: A causal theory. *Review of Philosophy and Psychology, 13*(4), 1097–1117.

Najenson, J. (2021). What have we learned about the engram?. *Synthese,* 199(3), 9581–9601.

Openshaw, J. (2023). (In defence of) preservationism and the previous awareness condition: What is a theory of remembering, anyway? *Philosophical Perspectives, 37:* 290–307.

Openshaw, J. (forthcoming). 'Generalised remembering', *Philosophical Studies*. DOI: 10.1007/s11098-026-02532-y

Openshaw, J. & Michaelian, K. (2024). Reference in remembering: towards a simulationist account. *Synthese, 203*(90). DOI: 10.1007/s11229-024-04508-z

Perrin, D. (2016). Asymmetries in subjective time. In *Seeing the Future: Theoretical Perspectives on Future-Oriented Mental Time Travel*. (Eds. K. Michaelian, S. B. Klein, and K. K. Szpunar). New York: Oxford University Press.

Perrin, D. (2021). Embodied Episodic Memory: a New Case for Causalism?. *Intellectica, 74*(1), 229–251.

Pessoa, L. (2022). *The entangled brain: How perception, cognition, and emotion are woven together*. Cambridge: MIT Press

Poldrack, R. A. (2010). Mapping mental function to brain structure: how can cognitive neuroimaging succeed?. *Perspectives on psychological science, 5*(6), 753–761.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the national academy of sciences, 98*(2), 676–682.

Ritchie, J. (2022). Naturalism as a stance. In De Caro, M., & Macarthur, D. (Eds.). (2022). *The Routledge Handbook of Liberal Naturalism* (pp. 190–202). Routledge.

Robins, S. K. (2020). Defending discontinuism, naturally. *Review of Philosophy and Psychology, 11*(2), 469–486.

Robins, S. K. (2022). Episodic memory is not for the future. In Sant'Anna, A., McCarroll, C. J., & Michaelian, K. (Eds.). *Current controversies in philosophy of memory* (pp. 166-184). New York, NY: Routledge.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of experimental psychology: Learning, Memory, and Cognition, 21*(4), 803.

Schacter, D. L. & Addis, D.R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transaction of the Royal Society B* 362: 773–786

Schacter, D. L., & Tulving, E. (1994). What are the memory systems of 1994? In D. L. Schacter & E. Tulving (Eds.). *Memory systems 1994*. The MIT Press.

Schwartz, A. (2020). Simulationism and the function (s) of episodic memory. *Review of Philosophy and Psychology, 11*(2), 487–505.

Schwartz, A. (2025). Are episodic memory and episodic simulation different in kind?. *Philosophical Psychology,* 1–24.

Sellars, W. S. (1962). Philosophy and the scientific image of man. In Colodny, R. (ed.), *Science, Perception, and Reality* (pp. 35–78). New York: Humanities Press/Ridgeview.

Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychological review*, 94(4), 439.

Slater, M. H. (2015). Natural kindness. *The British Journal for the Philosophy of Science*.

Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of learning and memory,* 82(3), 171–177

Suddendorf, T. & Corballis, M. (2007). The evolution of foresight: what is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences* 30: 299–313.

Sullivan, J. A. (2017). Coordinated pluralism as a means to facilitate integrative taxonomies of cognition. *Philosophical Explorations, 20*(2), 129–145.

Thomasson, A. (2023). How it All Hangs Together. In Miguel Garcia-Godinez (Ed.) *Thomasson on Ontology*. Palgrave Macmillan: 9–38.

Tulving, E. (1983). *Elements of Episodic Memory*. Oxford: Oxford University Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne, 26*(1), 1–12

Tulving, E. (2005). Episodic Memory and Autonoesis: Uniquely Human? In Terrace, H. S., & Metcalfe, J. (Eds.). *The missing link in cognition: Origins of self-reflective consciousness*. Oxford University Press.

Werning, M. (2020). Predicting the past from minimal traces: Episodic memory and its distinction from imagination and preservation. *Review of philosophy and psychology,* 11, 301–333.

Whewell, W. (1840/1847). *The philosophy of the inductive sciences* (2nd edition). John W. Parker.