Penultimate version. Forthcoming in Aronowitz, S. & Nadel, L. (eds.) *Memory, Space, and Time.* Oxford University Press. Please do not cite without permission.

A Place for the Memory Trace

Sarah Robins, Purdue University

skrobins@purdue.edu

Abstract

Memory traces are a persistent yet puzzling feature of our thinking about memory in both philosophy and memory science. In this chapter, I explore how further reflection on the explanatory role that memory traces play can benefit work across these disciplines. I divide my query into *philosophy-first* and *science-first* approaches to memory traces, identifying several distinct ways memory traces are understood and their comparative strengths and limitations. While none of the philosophy-first approaches to memory traces are successful, distinguishing amongst them may promote the development of further alternatives. The science-first approach highlights a way to generate such alternatives: from reflection on areas and eras of memory science where appeal to the trace concept is particularly useful. Here I focus on extracting an argument for memory traces from the recent resurgence of interest in the engram in neurobiology, arguing that it provides a novel and fruitful conception of the memory trace.

Keywords: *memory trace*, *engram*, *mental representation*, *mental image*, *causal theory of memory, episodic remembering*

1. Introduction

Aristotle characterized memory traces as impressions in wax.[1] Experiences are stamped on the mind "just as persons do who make an impression with a seal" (*de Memoria et Reminiscentia,* 450a). The successful formation, retention, and reanimation of such a trace requires that the mind's wax be in the right condition: just malleable enough. Too firm and it will fail to leave behind any impression at all. Too soft and it will be easily overridden by subsequent impressions.

Aristotle's depiction is, on the one hand, quaint. On the other, it shares important similarities with contemporary theorizing about memory. When pressed to define memory, appeals to representations remain a common feature. Take, for example, Yadin Dudai's (2007) definition of memory as "the retention over time of experience-dependent representations" (p. 15), or Morris

[1] Plato first proposed, but dismissed, the wax tablet view in the *Theatetus.*

Moscovitch's (2007) construal of memory as the "lasting internal representation of a past event or experience (or some aspect of it) that is reflected in thought or behavior" (p. 17).

Wax of course no longer works as a metaphor. Wax tablets were traded in for parchment and then for paper. Subsequent upgrades include the camera obscura, the phonographic record, the computer hard drive, the cloud.[2] While the metaphors have changed, the reliance on metaphor to build an analogy between memory and forms of information encoding, recording, and transfer remains. The memory trace thus remains as puzzling as it is persistent.

The lack of attention to the memory trace concept is shared amongst philosophers, psychologists, and neuroscientists—although the oversight manifests differently across disciplines. Philosophers are increasingly wary of memory traces and the need for them to account for memory. This is perhaps best evidenced by the recent emergence and growing popularity of *simulationist* accounts of memory, according to which memory is a way of imagining the past—a capacity whose reliability is not underwritten by traces (e.g., Michaelian 2021; Munro 2021). Even those who want to keep memory traces make a point of diluting the concept so as to render it more palatable. For example, some propose that traces are dispositional, rather than explicit (De Brigard 2020, this volume). Others pull back from the idea of localization, characterizing traces as distributed (Bernecker 2010). Still others cast the scope of traces wider, eschewing their dependence on the mind/brain and allowing traces to be embodied and extended (Sutton 2008). And still others keep the focus on traces as internal, but construe them as content free (Hutto & Peeters 2018) or minimal (Werning 2020). The range of proposals on offer reveals an underlying issue. While many who continue to think of memory traces as important, it is unclear whether there is any further consensus amongst them about what traces are, or the features of them that are critical to retain. The lack of internal agreement raises further questions about what is being rejected by simulationism and other purportedly non-trace accounts and whether the disagreements are substantive or semantic.

Amongst psychologists and neuroscientists, the commitment to memory traces remains more firmly in place, but there is little explicit attention given to what this entails. This is due in part to interest in exploring memory's role in a broader understanding of cognitive and neural dynamics – e.g., how memory is influenced by the brain's oscillatory dynamics in general and the production of spatiotemporal trajectories in particular (Buzsaki 2019), or accounts of episodic memory as a byproduct of a larger system for episodic simulation (Schacter & Addis 2007) or scene construction

[2] See Draaisma (2000) for an excellent overview of memory metaphors throughout history.

(Maguire & Mullally 2013). The neglect is also due to the concept's centrality; it is often simply 'baked in' to how memory is understood. Consider, for example, the book *Science of Memory: Concepts* (Roediger, Dudai, & Fitzpatrick 2007), which identifies 16 key concepts for memory science and features short reflections from 3-4 leading researchers on that concept. The list of concepts includes encoding, transfer, forgetting…but not the memory trace. Memory traces do, however, find their way into the book. They are the focus of entries on the introductory concept: memory. The quotations from Dudai and Moscovitch at the outset of this chapter were, in fact, drawn from those entries. This is of course only an anecdote, but it illustrates how the importance of a concept can coexist with its neglect. Memory, the capacity, involves retention of memories. A commitment at this level of generality is sufficient for extensive and fruitful investigation of the factors that influence what and whether remembering occurs and identification of the underlying cognitive and neural mechanisms. Such inquiry has produced many key insights about memory, but there is widespread acknowledgement that little progress has been made in understanding precisely what is stored and how (Maguire 2022; Poeppel & Idsardi 2022).

The aim of this chapter is to explore whether there is a place for the memory trace in our theorizing about and investigation of memory. To do so, I set aside these acknowledged differences in how philosophers and memory scientists engage with memory traces, exploring instead how reflection on the concept could be generally beneficial. I propose stepping back to evaluate the various explanatory roles for which memory traces are invoked in both philosophical and empirical contexts. Making these background, often implicit lines of reasoning explicit provides an opportunity to subject them to scrutiny. We can consider whether these explanatory roles are well-motivated and well-reasoned—and whether they in fact compel memory traces. We can also get a clearer sense of how memory traces are being understood, as distinct explanatory roles are likely to invoke distinct features. This provides a way of sorting amongst various philosophical views and their varying commitments. Whether a particular feature can be removed or diluted will depend, in part, on what role the trace is playing. It also makes evident the way that one's conception of the memory trace frames the empirical study of memory, shaping experimental design, investigation, and interpretation. Taking the time to make these guiding assumptions explicit and subject them to scrutiny can serve as an important reflection on experimental practice—perhaps suggesting reforms, but also identifying exciting new ways to move forward. To that end, the remaining discussion is guided by the following two questions:

1) What role are memory traces thought to play in an account of memory?
2) What must memory traces be like in order to play that role?

I divide exploration of these questions into *philosophy-first* and *science-first approaches*, covered in sections 2 and 3, respectively. In Section 2, I identify three standard lines of reasoning by which philosophers invoke memory traces, illustrating the distinct view of memory traces that comes from each. While none of the arguments is particularly successful in articulating a compelling and plausible account of memory traces, distinguishing amongst them may promote the development of further alternatives. The science-first approach in Section 3 highlights a way to identifying such alternatives: from reflection on areas and eras of memory science where appeal to the trace concept is particularly useful. Here I focus on extracting an argument for memory traces from the recent resurgence of interest in the engram in neurobiology, arguing that it provides a novel and fruitful conception of the memory trace.

2. Philosophy-First Arguments for Memory Traces

Below I survey three general lines of reasoning that philosophers have used to argue for the existence of memory traces.[3] Each subsection concludes with a brief evaluation. As will be shown, none of the available accounts offer a particularly compelling argument for the existence of memory traces. Each faces significant challenges, both in terms of the coherence of the explanatory role proposed and in terms of the empirical plausibility of traces that have the features such a role requires. The final subsection, 2.4, provides some reflections on this analysis and what it suggests as the way forward.

2.1 Representing the Past: Memory Traces as Mental Images

The first account of memory traces emerges from the need to explain our ability to represent past events or experiences. This route to traces can begin from a number of distinct philosophical commitments. One may, for example, subscribe to the view that all mental processes are representational, and so consider memory traces to be the particular form of representation required

[3] The arguments identified in this suggestion derive from my earlier survey of memory traces (Robins 2017). The current form involves some updates, most notably to ways of thinking about the features of memory traces corresponding to each explanatory role.

for memory. Augustine (2002) endorses a version of this view. This line of reasoning is best associated with the indirect realism of Locke (1975) and Hume (1978), according to which mental processes like perception and memory make us directly aware of representations and only indirectly aware of the features of the world that they represent. One could also have a more restricted sense of when mental representations are required and conclude that remembering requires memory traces, either because it has to be differentiated from perception (Martin 2015) or because remembering involves events or experiences that are no longer present or occurrent (Von Leyden 1961). Additionally, one could be led to this view of memory traces by considering the act of retrieval in particular. Retrieval involves selecting one past event from the set of such events retained in memory. The ability to do this requires an explanation and so the memory trace is invoked to explain how retrieval is possible (e.g., James 1890).

What must memory traces be like in order to play this role? While these arguments differ in scope, they converge on a view of memory traces as mental images: depictions of the prior events or experiences for which they are standing in. Often, these are conceived of as visual, but could be images based in other sensory formats—a song's chorus, the musky scent of your grandmother's perfume. If the trace were not a replica of the past event, then it could not fulfill its role as the object of thought or goal of retrieval. I can, for example, remember my 8th birthday party. The event itself is not available for me, as an adult, to perceive. A representation of the event stands in. Similarly, when attempting to remember this birthday party, I may have the experience of scanning back through a host of mental images of childhood parties and events. My search stops, and remembering occurs, once I find the image that depicts my 8th birthday party. Importantly, on this view, memory traces are personal level entities.[4] That is, they focus on the role of traces in the act of remembering as it is experienced by the rememberer herself. The memory trace is what a person is aware of when thinking about the past. The memory trace of my 8th birthday party is the mental image that I "see" in my mind, depicting the features of that experience I have retained.

The need to represent the past provides a role for memory traces that is easy to describe, but difficult to codify into an actual mechanism or process. In order for memory traces to play the role assigned to them, they must not only be mental images, but mental images that render their identity

[4] The intended contrast here is with the subpersonal level. Philosophers draw the personal/subpersonal distinction in subtly distinct ways (e.g., Drayson, 2014; Westfall, forthcoming; Dennett, 1969; Stich, 1978). The sense appealed to here is broadly compatible with the range of views on offer, emphasizing the difference between what is available from a first-person, introspective perspective and what is only available from a third-person perspective.

apparent. That is, a memory trace must not only depict the past experience, it must also make clear that what is being depicted 1) is of a particular experience (vs. some other experience) and 2) is a memory (vs. an imagination, hallucination, etc.). Each of these requirements poses a problem. Attempts to meet the first requirement make clear that the explanatory role of memory traces in retrieval is unstable. The argument is meant to show that retrieval requires a memory trace. But how is it that I know which trace to retrieve? That is: even if it 'looks like' my 8th birthday, how do I know what my 8th birthday looks like such that I can recognize it amongst images of other past events? If I know what it looks like already, then the initial motivation for the memory trace is undercut. A memory trace is not required for retrieval, as the argument initially supposed; the work is being done instead by knowledge I already possess. The other alternative thus seems preferable: it is not that I know in advance what I want to remember, but that I remember how to navigate to that item in my memory or recognize it upon presentation. But this route also leads to problems. The argument began from the claim that remembering requires a trace. In explaining how the trace is identified, appeal is made to another instance of remembering—remembering where this trace can be found or what features to look for. This remembering, too, must require a trace. A memory of how to locate or recognize a trace would be required for each trace. So the question can be raised again for this trace: how is the correct one identified? And so on. The need to account for each subsequent appeal to remembering presents the dilemma anew. Either this continues indefinitely, an infinite regress through which the traces required for a single instance of remembering pile up, or it stops with a form of remembering that does not require the retrieval of a trace – an abandonment of the central claim on which this began.[5]

Even if this line of argument were more successful, there is independent reason to be skeptical of the view of memory traces it entails. Consideration of what representing the past entails leads to a view of memory traces as mental images that are recognizable as depictions of past experiences. But there is significant reason to doubt the existence of any such identifying marker. Specific proposals are often subject to targeted criticism, as in Reid's (2002) challenge to Hume's (1978) claim that "vivacity" is what distinguishes memory images from those of perception, imagination, and the like. The more general concern derives from decades of empirical research. It is not simply that the feeling of remembering is hard to render precise, but that the various candidate markers are

[5] A version of this argument can be found in Heil (1978).

licentious—often accompanying inaccurate depictions of the past or representations of events that never occurred (e.g., Bernstein & Loftus 2009).

2.2 Sustaining Causation: Memory Traces as Filling the Gap

The second account of memory traces begins with reflection on a widely accepted claim: Remembering is a diachronic process. It involves an initial event or experience that is recalled during a subsequent experience. Reflection on what this claim involves yields the second account of memory traces: as the intermediaries required to explain the temporally extended process of remembering. By the middle of the 20th century, it was especially common to understand remembering's diachronic nature in causal terms—i.e., as seeing the subsequent recall of an event like one's 8th birthday party as an effect of the original event, the 8th birthday party itself.[6] From here, the commitment to memory traces falls out of an understanding of the nature of causation and what is required for sustaining a diachronic causal process.

There are many theories of causation available—and correspondingly, many particular ways to generate the memory trace commitment from one's view of causation. The commitment could arise from a general constraint on the nature of causation, specifically from the denial of action at a distance. Suspicion of remote causes is frequently found in discussions of causation, as in Leibniz's (2000) critique of Newton's invocation of gravity as an appeal to occult forces. Remembering is, in this respect, much like gravity: alleged to provide a causal influence on events from which it is spatiotemporally distant. Appeal to the memory trace as an intermediary precludes any accusations of the occult. On the broadly Humean view of causation, causes must precede their effects and immediately so. Otherwise, there is temporal gap by which the presumed or intended cause could be pre-empted. Memories may be effects of experience, but they rarely if ever follow directly from an experience. An intervening trace, formed as the result of experience and sustained until the time of remembering, thus provides the requisite contiguity (Bernecker 2008). The causal commitment to memory traces could also emerge from within a process view of causation (e.g., Salmon, 1984; Dowe, 1992), where causal relations are understood in terms of conserved quantities or mark transmission. On such a view, the memory trace would be the quantity conserved or the mark

[6] Commitment to this causal understanding of memory can be seen, for example, in Ayer (1956), Martin and Deutscher (1966), Shoemaker (1970), Anscombe (1981), and Armstrong (1987).

transmitted. Across these distinct views of causation, a core commitment emerges to the need for a memory trace to sustain the temporally-extended process of remembering.

To play this role, a memory trace must first be the kind of thing that can participate in causal relations—i.e., it must be physical.[7] It must also have features that support its intermediary, connective task. Cases of remembering can occur long after the experiences from which they derive, demanding the trace cover an extended temporal gap. The trace must therefore be the kind of physical thing that can persist across time. It must be stable, perhaps even static—capable of preserving the past experience's causal influence over extended periods, without disruption or deterrence from other influences.

The success of this argument for memory traces is dependent upon the success of these ways of thinking about the nature of causation. This is already suspicious. It would be preferable for the commitment to memory traces to derive from consideration of what remembering requires. Tying its fate to a particular way of thinking about causation opens up the possibility that the need for memory traces will dissolve if/when an alternative way of thinking about causation emerges. Indeed, this has happened. Concerns about action at a distance are less compelling given the kinds of things that are now thought to be capable of standing in causal relations. Moreover, there are counterfactual (Lewis 1973) and manipulability (Woodward 2003) theories of causation that lack contiguity, conservation, or transmission constraints. Such approaches to causation are increasingly popular, especially amongst those working in the biological and social sciences where processes like remembering are investigated.

There are additional problems with the memory traces themselves, as they are understood on this account. The memory trace earns its place by being stable and static, providing a past experience with a direct route toward expression in the present. But the brain in which such traces are presumed to reside seems an unlikely host. Neuroscientists are increasingly interested in thinking about the brain as a fluid system, better understood through dynamics and probabilities than stable structures and serial processes (e.g., Gerstner, Kistler, Naud, & Paninski, 2014). As Lynn Nadel (2007) has noted, this field-wide focus entails "demise for the fixed trace" (p. 181). There is, more generally, a lack of clarity in understanding how the contiguity requirement on remembering should be applied to the neural processes that underlie it. How close is close enough? One could, for

[7] Within this generic physicalist commitment, particular views of causation characterize causal relata differently, which could result in a number of distinct particular commitments: to the memory trace as an object, event, fact, conserved quantity, etc.

instance, worry about the insistence on spatiotemporal contiguity and the endorsement of the nearly consensus view of memory traces as synaptic. Synapses are, after all, *gaps*. It's hard to take this seriously as an objection; it seems both unfair and uninformed about neural systems. But even so it does make clear that which kinds of gaps in spatial and temporal relations are regarded as significant is likely to vary with both the systems under investigation and our interests in studying such systems. Contiguity itself does not provide much of a guide for identifying or understanding memory traces, or any other causes.

2.3 Remembering vs. Relearning: Memory Traces as Internal States

A final role for traces emerges from a tried-and-true philosophical method of conceptual analysis. Much of contemporary philosophy in the English-speaking world involves the analysis of linguistic predicates (i.e. concepts) and how to define them, or capture all and only their features or requirements.[8] Here the claim is than an analysis of *remembering* reveals memory traces to be a necessary feature, component, or requirement. Memory traces are simply built in to what it means to remember. The analysis proceeds via the *method of cases* (Machery 2017). One considers a range of scenarios, often hypothetical, where candidate features of the lacking or altered. Responses to the case—e.g., determination of whether the scenario presented involves remembering or not—are used to establish what the concept requires.[9]

Martin and Deutscher's (1966) *Remembering* is a primary example of this method, applied—as the title of their paper suggests—to remembering. Subsequent work by Bernecker (2010) and Debus (2010) add further cases to the analysis. Martin and Deutscher begin their analysis from the presumption that remembering requires, at a minimum, an accurate representation of something that has happened to the would-be rememberer. In order to remember my 8th birthday party, it must first be the case that I had such a party and also that my thinking about it now captures what happened.[10] They then go on to generate a range of hypothetical cases where both of these conditions are met and yet remembering does not occur. The key case that they consider is one of relearning: a case

[8] This chapter, with its focus on what memory traces could be, is hardly an exception.
[9] Whether responses to thought experiments should be understood as intuitions, reasoned judgements, etc. is a matter of current debate in philosophy. Nothing in the brief characterization here should be taken as favoring any particular view on this issue.
[10] Martin and Deutscher do not provide much further elaboration on what accuracy requires. There are a host of further questions one can press about this requirement (e.g., are omissions and commissions equally problematic? How much must be represented in order to achieve accuracy?), but I am setting them aside for present purposes.

where a person learns something, forgets it, and then reacquires it from another source. Their example involves a person named Kent, who is in an accident and later suffers amnesia that causes him to forget the accident (and other events). Kent is then provided with information about the event from another source, allowing him to re-represent it. In such a case, it is proposed, Kent has *relearned*, not *remembered,* the event.[11] To account for the difference between remembering and relearning, remembering must have a feature that is missing or malformed in the relearning case. The feature that is missing in cases of relearning is the memory trace. In cases of relearning, the memory trace has been lost. In cases of remembering, it has been retained.

So what would a memory trace have to be like in order to support this distinction? The case of relearning is designed so that it matches remembering on a number of features. Kent's case is meant to be like remembering in the existence of the remembered event (the accident), the accuracy of the current thoughts about that event—and even in the causal connection between the past event and the current thoughts about it.[12] The key difference between the two cases is the relation of that causal chain to the would-be rememberer. In remembering, but not in relearning, the causal chain stays *inside* the person, in some intuitive, but difficult to articulate sense. The distinction is best understood, I have argued (Robins 2016) as a *cognitive* difference. It is an assessment of the causal history of the current thought about the past event. Is the capacity to produce it something that has been retained or something that has been regained? To be remembering, that capacity must reside inside the cognitive system that produces the thought being assessed.

This of course gives rise to questions about the mind's perimeter. The cognitive boundary has been challenged from many directions in recent decades, as philosophers and cognitive scientists have argued that the mind extends out into the world and other people, as well as down into various parts of the body. Luckily, differentiating remembering from relearning does not require settling this issue. Wherever that boundary is drawn, remembering will fall inside it and relearning outside. Memory traces are internal states that produce instances of remembering.

The biggest challenge to this argument for memory traces is the number of philosophers who are not convinced by consideration of relearning. The case of Kent, for some, fails to elicit the

[11] When psychologists talk about relearning, they are often referring to a slightly different phenomenon: the re-exposure to information when one has not previously forgotten that information. The "savings during relearning" paradigm can be traced to Ebbinghaus (see Nelson, 1985).

[12] In this way, Martin and Deutscher's analysis of remembering includes commitment to a causal condition on remembering. It could be possible for someone to develop an alternative analysis that advocates for distinguishing remembering from relearning without the causal condition.

presumed response. Kent is remembering—or at least the possibility that Kent is remembering is not ruled out by the conditions as stated by Martin and Deutscher (Michaelian 2016). It is further unclear whether the distinction between remembering and relearning matters beyond contrived philosophical contexts. There are some instances where it seems to matter: the distinction between testimony and hearsay, or a perfect exam that results from studying versus a cheat sheet. But in many cases, even if the difference is understood, it is impossible to tell. As Martin and Deutscher (1966) themselves note, many recollections from early childhood are like this. We often do not know whether events from our early years are ones we remember or ones we remember the retelling of from others.

Even if this line of argument in favor of memory traces were particularly compelling, it does not provide an especially rich account of memory traces. It tells us only that they are states internal to the cognitive system, derived from past experiences.

2.4 **So What Now?**

Above I identified three distinct lines of argument that lead to a commitment to memory traces, each yielding a distinct conception of what traces are like: the memory trace as mental image, the memory trace as stable causal force, and the memory trace as internally maintained state. None of the features identified were in and of themselves surprising: representational, causal, and historical features are familiar to accounts of memory traces, as the definitions from the introduction illustrate. What is interesting, however, is how the features align with particular arguments. While these features are often lumped together in discussions both for and against traces, the above analysis makes clear how distinct features are connected to distinct ways of arguing for traces.

For example, the representational, imagistic features of memory traces are essential for accounts concerned with representing the past. Such views need not have any additional requirement that memory traces be internal or supported by the right kind of causal linkages. Representational views could even deny that remembering is causal. Conversely, views that derive from a commitment to remembering as causal one could eschew any commitment to representational or imagistic features and characterize traces as merely sustaining a causal connection (e.g., Werning 2020). Even the view of memory traces that derives from considering cases of relearning can be formulated without the features that stem from the other two lines of argument. The internal stages need not support mental imagery. And while it is true that Martin and

Deutscher's (1966) analysis does presuppose that remembering is causal, this need not lead to a requirement that that traces be static and stable. They could simply endorse one of the available views of causation that does not preclude spatiotemporal gaps.

The above evaluation demonstrates the importance of excavating the line of reasoning being used to defend the idea of memory traces, as it is critical for determining which of the myriad features often associated with memory traces are needed for the argument at hand. Unfortunately, it also demonstrates something else: none of these lines of argument is particularly successful. Each is ineffective at securing a clear and compelling explanatory role for memory traces. And, in each case, there are also independent reasons to be suspicious about the existence of memory traces with the requisite features. In searching for a place for the memory trace, it seems that it would be more profitable to go in a new direction.

3. Science-First Arguments for Memory Traces

In this section, I turn to what I loosely describe as science-first ways of arguing for memory traces. As stated in the introduction, the memory trace often serves as a background commitment for empirical investigation of memory mechanisms. There are times, however, when the concept receives more explicit attention. Although the focus in such times is not generally on arguing for memory traces, such an argument may be identifiable from the way in which the interest in and usefulness of the trace emerges.

To this end, I am particularly interested in the resurgence of work on the engram in contemporary neurobiology. Elsewhere I have argued that this line of research marks a significant change in the kinds of inquiry into memory mechanisms that neurobiologists of memory are conducting (Robins 2018; Robins 2023). Here I focus on the motivation for that change and how it might be built up into an argument for memory traces. In 3.1, I discuss engram research and its significance in the current context. In 3.2, I identify the explanandum and how it might be developed into an argument for memory traces. In 3.3, I assess its explanatory adequacy in comparison to other approaches.

3.1 The Engram Renaissance

The neurobiology of memory is currently experiencing an "engram renaissance"' (Josselyn, Köhler, & Frankland, 2017: 4647). Researchers working in this area do not regularly refer to their

work in this way, and indeed the use of this phrase in the paper just cited is something of an offhand remark. I have amplified it here because it is apt for characterizing the significant change the field is undergoing. As the term *renaissance* implies, the current fervor does not involve a new idea or concept, but the rebirth of an old one. "Engram" was coined by the zoologist Richard Semon (1904/1921) more than a century ago. The term has been in use since, but only intermittently. A *Web of Science* search shows that Semon's book was cited more in the decade from 2010-2020 than in the previous ten and half decades since its initial publication.[13] *Renaissance* also implies a time of new discovery and the exploration of themes and connections that disrupt established boundaries of inquiry. In this way, research into the engram has been exciting not only for what it reveals about the basic mechanisms of memory, but for the opportunities it provides to connect with broader areas of memory science: investigations of false memory in cognitive psychology, models of long-term memory consolidation in neuropsychology, and the treatment of Alzheimer's.

To make clear how the engram renaissance is relevant to this investigation of memory traces requires saying more about both the engram and recent discoveries. First, the engram. Semon (1904/1921) defined the engram as the physical or chemical change to the brain that results from learning. Semon coined the term to guide inquiry, not to label a discovery. In this way, the engram is a more technical, scientific version of the memory trace. How much overlap there is between the two terms depends on how widely one's scope of inquiry is drawn. As we saw above in Section 2, the memory trace can be understood in many ways, not all of which invoke physical features, much less the brain. For philosophers, then, the two concepts can diverge significantly. Amongst memory scientists who are focused on cognitive and neural systems, however, the two are generally treated as synonyms. Insofar as the engram is a way of conceiving of memory traces, here I will follow the scientific practice of treating "memory trace" and "engram" as interchangeable.

In the neurobiology of memory, the engram has long persisted as something of a regulatory ideal: a commitment that guides inquiry but is not taken too literally. That is, the investigation of the mechanisms of memory is framed around the presumption that memory involves the encoding, storage, and retrieval of engrams, but studies are designed around the general investigation of these processes, not the search for individual engrams. With the development of new experimental tools, this has changed. Advances in activity-dependent cell-labeling (Mayford 2014), in vivo calcium imaging (Yang and Yuste 2017) and chemogenetics (Armbruster et al. 2007) have brought increased

[13] 83 vs. 29. Search conducted January 2021.

precision to the investigation and manipulation of memory mechanisms. But the real driver of change has been *optogenetics*, a technique for controlling neural activity with light (Deisseroth 2010). Optogenetics allows neuroscientists to target a selective group of neurons and activate or inhibit their expression in living, behaving organisms.

In application to memory, this has meant the ability to create light-responsive engrams (Liu et al. 2012). Working mostly with mice, researchers have shown that optogenetics can be coupled with standard spatial memory paradigms so that the neurons active during the encoding of a spatial memory will subsequently express a light-sensitive protein. The expression of this protein makes it possible to reactivate the neurons that encoded the memory by exposing them to light. Memory retrieval via light switch.[14] The availability of this new form of intervention affords researchers a host of opportunities to explore the features of memory formation, storage, and reactivation under a host of conditions. This has led to a range of fascinating discoveries, including: the production of false memory in non-human animals (Ramirez et al. 2013); the ability to change the valence of a memory from positive to negative and vice versa (Redondo et al., 2014); recovery of memory in early-stage Alzheimer's models (Roy et al. 2016); and the creation and implantation of an artificial engram (Vetere et al. 2019).

My interest here is not in a direct assessment of this research program, nor the validity or significance of any of these results. Regardless of how one views this line of research, its growth and prominence is undeniable. My interest is in the role of the engram/memory trace within this research program. Why is this new approach to investigating memory re-exciting talk of the engram? It cannot be written off as simply the result of a new investigative tool or discovery. After all, there have been many such tools over time. Viewed from a wider lens, the broad consensus that now exists about the mechanisms of memory is the result of a steady stream of tool development and discovery across several decades and areas of neuroscience (Craver 2003; Silva, Landreth, & Bickle, 2014). My question is: why is the engram concept of such interest, *now* as opposed to other significant moments in the history of memory investigation? Why did it not occur alongside the development of fMRI or CRISPR?

I want to suggest that the answer to this question has to do with the specific kind of change that the optogenetic investigation has brought about. There has been a shift, from investigating *mechanisms* to investigating *memories*. The neuroscience of memory has long been a study of memory's

[14] A more thorough review of the method can be found in Robins (2023).

mechanisms (Craver 2007). How are long-term memories formed? What is the nature of consolidation (and reconsolidation)? These questions are answered by identifying neural mechanisms, organized clumps of entities and activities, and characterizing how their interactions change under a range of conditions, perturbations, and interventions. Now, with the advent of optogenetics, researchers can use this general mechanistic understanding to explore the formation, persistence, and modification of particular memories. That is, optogenetics allows for the identification, investigation, and manipulation of token activities in the memory system.

This provides an answer to the question with which this section began: the memory trace/engram becomes relevant and interesting when researchers have the capacity to identify, isolate, and manipulate memories of particular past experiences. As the available tools make it possible to study the retention of specific, token occurrences, the idea of the engram returns. Can this be used to develop a new line of argument for memory traces—an argument built around their role in explaining the retention of particular past experiences? I explore this in the next subsection below.

3.2 Retaining Particulars

First, we must determine whether the retention of particulars is a phenomenon worth explaining. If it rarely occurred, or occurred only in limited or contrived contexts, or was marginal in some other way, then this would not be a particularly promising line of inquiry.

We do in fact remember particular past encounters—at least sometimes. I will substantiate this claim with evidence below, but it's truth should (I hope) seem fairly clear to most readers simply from experience. It occurs in cases of one-shot learning and instances of episodic remembering. Of course, not all learning happens this way and not all attempts at episodic remembering are successful. What matters for present purposes is that it does happen. The scope of 'we' in this claim is intended to be vague and somewhat broad. I have in mind not only humans, but also other animals whose memories are well-studied (e.g., mice, scrub jays), possibly more. I do, however, take 'remember' to be something of a success term here. The intention is to highlight cases where we get it right, without worrying (at least for now) about how to set specific standards of accuracy and correctness. 'Particular past encounters' is a mouthful, but the aim of such a term is to remain neutral amongst memory taxonomies and ways of defining episodic or event memory while still highlighting token occurrences. How tokening occurs, and how the boundaries of tokens are set is

an interesting puzzle, but not one that needs to be settled for establishing this basic claim about retaining particulars. The spatiotemporal boundaries of token occurrences are likely to be set by mechanisms of event segmentation, which may vary across species, individuals, and/or contexts.

The retention of particular past experiences is something that a theory of remembering ought to be able to explain. Retaining specific facts, ideas, and experiences is a quintessential form of memory. When asked to think of a memory, such occurrences are often what's selected. It is an emblematic form of memory, arguably where its most distinct from other capacities like perception and belief. Appeal to a memory, or engram, provides a straightforward explanation of it: features of the occurrence are encoded, retained, and later reactivated. Whether such an explanation is compelling enough to support an argument in favor of memory traces depends on whether there are other ways of explaining the retention of particulars and, if so, their comparative strength. That is, the appeal to memory traces to explain the retention of token occurrences could amount to a successful argument for the existence of memory traces if it can be shown to be the best possible explanation of this phenomenon.[15]

Successful retention of particulars has not received a lot of direct attention in the past several decades. Instead, explanatory frameworks have been centered around 1) memory errors and 2) the role of patterns, generalizations, and abstractions in remembering processes, both as a way of accounting for 1) and in attempt to align memory with theories of cognitive architecture more generally.

Approaches to memory mechanisms that highlight their role in broader cognitive and neural processes, like those identified in the introduction, are prime examples. Such views are at pains to explain how memory errors can occur as frequently and systematically as they do. Dan Schachter, a leading proponent of the episodic simulation hypothesis, describes how evidence of false memory has shaped his thinking about the nature and function of memory more generally, saying that "it makes little sense that evolution would have yielded such a deeply flawed system" (2019, p. 265). Instead, he and others have come to support an alternative way of looking at memory systems, such that memory errors are viewed "as byproducts of otherwise adaptive features of memory" (Schacter, 2019: p. 265) Similarly, philosophers have recommended reorienting our thinking about the relation

[15] De Brigard (2020) offers a similar Inference to the Best Explanation argument for memory traces, where what's being explained is the causal connection between an experience and its subsequent recollection.

between memories failures and successes such that they can both be understood as cases of the memory system "doing what it is supposed to do" (De Brigard 2014, p. 172).

What is it that memory is supposed to be doing? What are its adaptive features? There are a range of distinct positive proposals on offer. They all supplant the standard approach to memory as a faculty for retention, and with it the encoding-storage-retrieval approach through which memory traces may play a role. They highlight instead a broader cognitive aim—e.g., prediction, inference, decision-making, narrative or schematic cohesion—and posit a more generalized and distributed form of information storage suited to this aim. Addis (2020), for example, thinks of constructive episodic simulation as a critical cognitive capacity. Constructive episodic simulation is here understood as the ability to generate representations of what the world is like from a first-person (episodic) perspective across temporal and conceptual dimensions. I can construct episodic simulations about what I will do tomorrow, or next month, or 10 years from now, or in retirement. I can also construct episodic simulations of what I should have done yesterday, or what I would have done if circumstances had been otherwise, or what I would do now with difference resources, or what I could do if anything were possible. To support these simulations, the underlying cognitive systems is organized to support schematic representations and associations between concepts and other event components. Schemas provide general, thematic structures for events and activities and associationist networks provide further information about things and features that tend to co-occur. Together, these can be used to build compelling episodic simulations of events—ones that should happen, might have happened, or could happen in the future.

Remembering what did happen in the past is, on this view, yet another particular application of this general capacity. Event schemas and associated features are used to build a plausible representation of the past. This works reasonably well. When the events we are trying to remember fit with these schemas and associations, the resultant simulations get it right, or mostly so. But it also explains how memory errors occur, as the influence of schemas and associations may lead to the insertion of themes and elements that were not a part of the past event in question. While the influence on remembering is bothersome, it is useful for other forms of episodic simulation, and so understandable as a feature of this system.

Addis' (2018; 2020) view is only one example, but it is reflective of a general explanatory strategy that de-emphasizes traces. No one denies that successful remembering occurs. And gestures are often made to suggest how such cases could be explained in terms of these overall frameworks. Still,

the focus is elsewhere—on broader cognitive tasks and the architectural principles by which cognitive systems achieve them. By highlighting the retention of particulars, work on the engram in neurobiology offers an opportunity to direct some of our attention back to this phenomenon and whether it is explained well-enough on such approaches. I take this up in the next section.

3.3 Memory Traces and Explanatory Adequacy

Some views of memory are generally focused on accounting for memory errors and/or memory as a sub-component of more general cognitive processes. On such approaches, successful retention of particulars is not denied, it is just not given much attention. The lack of attention has consequences, which challenge the explanatory adequacy of such proposals.

This is well-illustrated in a recent study by Diamond and colleagues (2020) that explores the accuracy of event memory in real-world contexts. They measured memory for two "complex and immersive yet verifiable real-world experiences" (p. 1545): 1) a training on mask-fitting procedure given to hospital employees, and 2) a tour of a hospital foyer with lots of distinctive art and architecture. Importantly, participants were recruited to the study *after* their respective experience, removing concerns about experimental context and participant expectations from the encoding process. Once participants were recruited to the study, the researchers used free recall techniques to elicit memories and measure how much of the initial information had been retained, at intervals ranging from two days (for the hospital tour) to two years (for the mask-fitting training). Across both experiences, they found that participants were able to provide many event-specific details in their recollections—and, importantly, the details provided were highly accurate. For example, participants who had experienced the mask-fitting training were able to generate an average of 40-50 details about the experience, more than 95% of which were confirmed as accurate, even when being tested 1-2 years after it occurred.

Alongside these measures, Diamond and colleagues (2020) conducted a survey of memory scientists and other academics, where they presented survey participants with hypothetical cases modeled on the real-world experiences assessed above, asking them to estimate how accurate a person's recall would be in such a scenario. Survey responses showed low confidence in people's capacity for accurate retention. When asked, for instance, to estimate how well a healthy 30-year-old would perform when asked to recall the details of a museum tour 2 days after the experience, the

researchers surveyed estimated that their accuracy would be between 20-30%. For the hospital tour participants in this age range, accuracy was around 94%.

The Diamond et al. (2020) study demonstrates that successful retention of particular past experiences occurs and illustrates an ingenious way of uncovering it. The study simultaneously shows the problems that can arise when theoretical and empirical inquiry is too focused on memory errors. As they state, the results show that many researcher's "views of human memory accuracy are overly pessimistic" (2020, p. 1552). This finding matters not only because of the disconnect it reveals between everyday remembering phenomena and researchers' suppositions about it, but also because these expectations are ones that researchers are unlikely to correct if they are not motivated to investigate. One's estimate of how likely successful retention is will shape one's expectation of whether it is fruitful to explore empirically or incorporate into one's theoretical framework.

With these expectations of memory's accuracy challenged, we are then free to explore and notice other domains where retention of particulars occurs. This could include cases of highly-superior autobiographical memory (LePort et al. 2017) or broader investigations of expertise, both memory-specific (Foer 2012) and more generally (Ericsson and Lehman 1996). We might also find cases in less obvious places, like models of PTSD (Zhang et al 2020) and addiction (Chiew & Adcock 2019) where the retention of particulars has less positive consequences.

This may also, in turn, lead to critical reflection on the explanatory adequacy of false memory frameworks for the very phenomena they were developed to explain. Eyewitness testimony has, for example, been one of the key forms of false memory to generate attention and concern. Evidence of the ease and extent to which firsthand reports of past experiences can be manipulated and distorted has challenged how many think about the significance of testimony, challenging its long-standing pride of place in legal contexts (Loftus 2003 for review). While the effects of interrogation techniques on testimony are clear and well-documented, their overall prevalence and significance may have been overstated. As Wixted and colleagues (2018) argue, our understanding of how errors can come about in testimony is consistent with a view that such testimony can be accurate when such forces are kept at bay.

How troubled one is by cases that do not fit the general pattern of explanation critically depends on how many cases one thinks there are. If there are only a handful of such cases, then even the realization that they are not captured well, or at all, by one's framework may not be particularly concerning and could possibly be absorbed with little to no alteration to one's theoretical

commitments. A view of remembering as supported by a cognitive system of schemas and associations, like Addis' (2020) view above, could likely take on board a few non-schematic, individual event representations. As more such cases need to be accommodated, this becomes more difficult and continuing to absorb the anomalies becomes inadvisable. Taking on too many exceptions risks negating the explanatory virtues of the proposed architecture. Appeal to schemas is meant to provide an account of how information can be efficiently and effectively organized. As the amount of non-schematically organized information grows, the ability for schemas to play their organizational, streamlining role is complicated. So too for systems built around principles of association, frequency, narrative cohesion, etc. The power and explanatory benefit of such organizational frameworks comes from their breadth. This is not, of course, to say that schemas have no role to play in scaffolding particular memories. It seems plausible that they play a key role in preserving content, facilitating pattern completion, etc.[16]

This criticism should not be understood as a recommendation that such cognitive frameworks be rejected. Schematic, associationist, and frequentist processes clearly play a role in many cognitive operations. The question, as I see it, is how to understand the scope of their explanatory power. Should they be used to explain all processes and operations? If we identify phenomena/systems/processes that do not fit well into these frameworks, should we attempt to subsume them or make space for distinct, complementary architectures? My preference is to support the latter. Setting constraints on the applicability of these pattern-based processes would be a way to retain their explanatory virtues, while allowing for alternative explanation of phenomena that are ill-suited to that framework. What the details of such a proposal (likely proposals) would look like is still far from clear. But it's still progress to recognize the work that needs to be done.

To summarize this section, I have used the case of the engram renaissance in neurobiology to argue that the memory trace is a useful concept for memory scientists when they are exploring the retention of particular past experiences. In reflecting on this phenomenon of retaining particulars, we can come to see that it is an important, central phenomenon to be explained. As the situation stands currently, trace-based explanations of our ability to retain particular past occurrences fare better than alternatives focused on the role of more general cognitive processes. It is possible that this apparent advantage is illusory. As noted above, the need to explain retention of particulars does not receive much direct attention. It could be that, once they do so, more plausible and competitive

[16] Thanks to Sara Aronowitz for pressing this point.

explanations will emerge. For now, there is at least a place for the memory trace worth investigating further.

3.4 Memory Traces as Discrete Entities

The last section sketched an explanatory role for memory traces: as the best available explanation of our ability to retain particular past occurrences. We can now go on to ask the second question: what must memory traces be like in order to play this role? Here I elaborate, briefly, on the features such a view of memory traces compels.

To account for the ability to retain particular past occurrences, memory traces must be discrete entities. Discrete here stands in opposition to continuous. The form of retention memory traces support must be one that allows for particular past encounters to be retained in ways that can, at least sometimes, survive interference from retention of other particular past encounters and more general forms of information storage, processing, and updating. In other words, if retaining information from a particular experience is explained by a trace, then the trace that makes that possible has to be built such that it is capable of remaining distinct from other traces and from other non-trace kinds of information.

This commitment to discreteness places a constraint on both the structure and content of the memory trace. Discreteness is, in fact, a structural commitment in the sense that Hebb (1949) first proposed as a constraint on the engram, as an intended contrast with diffuse and dynamic patterns of activity. The claim that memory traces are *discrete* should not be mistaken for the claim that memory traces are *local*. Memory traces could be distributed, in a sense operative at multiple levels of organization: across networks, regions of the brain, neural populations. Similarly, traces can be *discrete* without being *static*. The kind of retention they support must be capable of resilience amid other instances and form of retention, but the precise vehicles by which that resilience is supported could change over time. In order for discrete traces to explain the retention of particular past encounters, they must not only contribute causally to their remembering but do so in a way that is reflective of the content acquired or encoded from that encounter. They must, in other words, be information-bearing or representation-supporting in some critical way. How this content is to be understood is not fully clear. There is likely room for a range of proposals as to what this content is and how it is carried by the trace.

Thus, even with this commitment to the memory trace as a discrete entity, there are a range of ways in which the constraints on its structure and content could be understood. And from here, even more work to be done articulating how their structure and content are related to one another—both at any particular moment, and over time. Exploring various proposals would offer an increasingly rich understanding of what memory traces are or could be, perhaps supporting multiple distinct forms or competing proposals, the testing of which could yield fruitful new lines of theoretical and empirical inquiry.

4 **Conclusion**

The memory trace has a long and complicated history. Whether interest in the memory trace should be restricted to retrospective storytelling is, for the moment, an open question. Many would happily leave it behind as memory research goes forward. The recent development of optogenetic tools has led at least some memory researchers to consider the concept worth dusting off and carrying forward. The role they see for memory traces, in the retention of particular past encounters, is a promising one—and one that is otherwise neglected in current theorizing. How far such a proposal can go remains to be seen. Regardless of the outcome, exploring whether there is a place for the memory trace provides a critical moment of reflection on our ways of thinking about the nature and purpose of remembering.

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,* 64-88.

Addis, D. R. (2020). Mental time travel? A neurocognitive model of event simulation. *Review of Philosophy and Psychology*, *11*(2), 233–259.

Anscombe, G.E.M. (1981). Memory, experience, and causation. In G.E.M. Anscombe (Ed.) *Collected Philosophical Papers, Vol. II: Metaphysics and the Philosophy of Mind* (pp. 120–130). Oxford, UK: Oxford University Press.

Aristotle. *De Memoria et Reminiscentia* (W. D. Ross, trans, 1955). Oxford, UK: Clarendon Press.

Armbruster, B.N., Li, X., Pausch, M.H., Herlitze, S., & Roth, B.L. (2007). Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proceedings of the National Academy of Sciences, 104,* 5163-5168.

Armstrong, D. M. (1987). Mental concepts: the causal analysis. In R.L. Gregory (ed.), *The Oxford Companion to the Mind* (pp. 464–465). Oxford: Oxford University Press.

Augustine of Hippo (2002). *On the Trinity: Books 8-15*. Trans. McKenna, S. Ed. Matthews, G.B. Cambridge: Cambridge University Press.

Ayer, A.J. (1956). *The Problem of Knowledge*. Harmondsworth, UK: Penguin.

Bernecker, S. (2008). *The Metaphysics of Memory*. Springer.

Bernecker, S. (2010). *Memory: A Philosophical Study*. Oxford University Press.

Bernstein, D.M. & Loftus, E.F. (2009). How to tell if a particular memory is true or false. *Perspectives on Psychological Science, 4,* 370-374.

Buszáki, G. (2019). *The Brain from Inside Out.* Oxford University Press.

Chiew, K. & Adcock, A. (2019). Motivated memory. In K.A. Renninger & S.E. Hidi (eds) *The Cambridge Handbook of Motivation and Learning* (pp. 517-546). Cambridge: Cambridge University Press.

Craver, C. F. (2003). Interlevel experiments and multilevel mechanisms in the neuroscience of memory. *Philosophy of Science*, 69, S83–97.

Craver, C.F. (2007). *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford University Press.

De Brigard, F. (2013). Is memory for remembering? Recollection as a form of episodic hypothetical thinking. *Synthese*, *191*(2), 155–185.

De Brigard, F. (2020). The explanatory indispensability of memory traces. *The Harvard Review of Philosophy, 27,* 23-47.

De Brigard, (this volume). Simulationism and Memory Traces. In S. Aronowitz & L. Nadel (eds), *Space, Time, and Memory*. Oxford University Press.

Debus, D. (2010). Accounting for epistemic relevance: a new problem for the causal theory of. Memory. *American Philosophical Quarterly, 47,* 17-29.

Deisseroth, K. (2010). Optogenetics: Controlling the Brain with Light. *Scientific American*.

Dennett, D.C. (1969). *Content and Consciousness*. Cambridge, MA: MIT Press.

Diamond, N.B., Armson, M.J., & Levine, B. (2020). The truth is out there: accuracy of recall of verifiable real-world events. *Psychological Science, 31*, 1544-1556.

Dowe, P. (1992). Process causality and asymmetry. *Erkenntnis, 37*, 179-196.

Draaisma, D. (2000). *Metaphors of Memory: a history of ideas about the mind*. Cambridge, UK: Cambridge University Press.

Drayzon, Z. (2014). The personal/subpersonal distinction. *Philosophy Compass, 9,* 338-346.

Dudai, Y. (2007). Memory: it's all about representations. In Science of Memory Concepts, ed. H. L. Roediger, Y. Dudai, and S. M. Fitzpatrick, 13-16. Oxford: Oxford University Press.

Ericsson, K.A., & Lehman, A.C. (1996). Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47,* 273-305.

Foer, J. 2011. *Moonwalking with Einstein*. New York: Penguin Press.

Gerstner, W., Kistler, W. M., Naud, R. & Paninski, L. (2014). *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition.* Cambridge: Cambridge University Press.

Hebb, D. (1949). *The Organization of Behavior.* New York: Wiley & Sons.

Heil, J. (1978). Traces of things past. *Philosophy of Science, 45*, 60–72.

Hume, D. (1978). *A Treatise of Human Nature,* ed. L.A. Selby-Bigge, 2nd ed., Oxford: Clarendon Press.

Hutto, D. & Peeters, A. (2018). The Roots of Remembering. In K. Michaelian, D. Debus, & D. Perrin (eds.) *New Directions in the Philosophy of Memory*. Routledge.

James, W. (1890). *The Principles of Psychology*. London, UK: Macmillan.

Josselyn, Sheena A., Köhler, Stefan, and Frankland, Paul W. (2017). Heroes of the Engram. *Journal of Neuroscience, 37,* 4647-4657.

Leibniz, G.W. *G. W. Leibniz and Samuel Clarke: Correspondence*; Hackett.
Leport, A.K.R., Stark, S.M., McGaugh, J.L., & Stark, C.E.L. (2017). A cognitive assessment of Highly Superior Autobiographical Memory. *Memory, 25,* 276-288.

Lewis, D. (1973). Causation. *Journal of Philosophy, 70,* 556-567.

Liu X., Ramirez, S., Pang, P., Puryear, C., Govindarajan, A., Deisseroth, K., & Tonegawa S. (2012). Optogenetic stimulation of a hippocampal engram activates fear memory recall. *Nature,* 484, 381–385.

Locke, J. (1975). *An Essay Concerning Human Understanding*. P.H. Nidditch (ed), Oxford University Press.

Loftus, E. F. (2003). Our changeable memories: Legal and practical implications. *Nature Reviews: Neuroscience*, *4*(3), 231–234.

Machery, E. (2017). *Philosophy Within its Proper Bounds*. Oxford University Press.

Maguire, E. (2022). Does memory research have a realistic future? *Trends in Cognitive Science, 26*, 1043-1046.

Maguire, E.A., & Mullally, S.L. (2013). The hippocampus: A manifesto for change. *Journal of Experimental Psychology: General, 142,* 1180-1189.

Martin, M.G.F. (2001). Out of the past: episodic recall as retained acquaintance. In C. Hoerl & T. McCormack (eds.) *Time and Memory: Issues in Philosophy and Psychology* (pp. 257-284). Oxford University Press.

Martin, C.B., & Deutscher, M. (1966). Remembering. *The Philosophical Review, 75,* 161-196.

Mayford, M. (2014). The search for a hippocampal engram. Philosophical Transactions of the Royal Society B: Biological Sciences, 369, 20130161.

Michaelian, K. (2016). *Mental time travel: Episodic memory and our knowledge of the personal past*. MIT Press.

Michaelian, K. (2021). Imagining the past reliably and unreliably: Towards a virtue theory of memory. *Synthese*, *199*(3–4), 7477–7507.

Moscovitch, M. (2007). Memory: why the engram is elusive. In Science of Memory Concepts, ed. H. L. Roediger, Y. Dudai, and S. M. Fitzpatrick, 17–22. Oxford: Oxford University Press.

Munro, D. (2021). Imagining the actual. *Philosophers' Imprint, 21,* 17.

Nadel, L. 2007. "Consolidation: The Demise of the Fixed Trace." In Science of Memory Concepts, ed. H. L. Roediger, Y. Dudai, and S. M. Fitzpatrick, 177–82. Oxford: Oxford University Press.

Nelson, T.O. (1985). Ebbinghaus' contribution to the measurement of retention: savings during relearning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 472-479.

Poeppel, D. & Idsardi, W. (2022). We don't know how the brain stores anything, let alone words. *Trends in Cognitive Science, 26,* 1054-1055.

Ramirez, S., Liu, X, Lin, P., Suh, J., Pignatelli, M., Redondo, R.L., Ryan, T.J., Tonegawa, S. (2013). Creating a false memory in the hippocampus. *Science*, *341*, 388–391.

Redondo, R.L., Kim, J., Arons, A.L., Ramirez, S., Liu, X., & Tonegawa, S. (2014). Bidirectional switch of the valence associated with a hippocampal contextual memory engram. *Nature, 513*, 426–430.

Reid, T. (2002). *Essays on the Intellectual Powers of Man,* D.R. Brooks (ed.). Pennsylvania State Press.

Robins, S.K. (2016). Contiguity and the Causal Theory of Memory. *Canadian Journal of Philosophy, 47*, 1–19.

Robins, Sarah K. 2017. "Memory Traces." In Routledge Handbook of the Philosophy of Memory, ed. S. Bernecker and K. Michaelian, 76–87. London: Routledge.

Robins, S. K. (2018). Memory and optogenetic intervention: separating the engram from the ecphory. *Philosophy of Science*, *85*(5), 1078–1089.

Robins, S. (2023). The 21st century engram. WIREs Cognitive Science, e1653. https://doi.org/10.1002/wcs.1653

Roediger, H.L., Dudai, Y. & Fitzpatrick, S. (eds) (2007). *Science of Memory: Concepts*. Oxford University Press.

Roy, D. S., Arons, A., Mitchell, T. I., Pignatelli, M., Ryan, T. J., & Tonegawa, S. (2016). Memory retrieval by activating engram cells in mouse models of early Alzheimer's disease. Nature, 531, 508–512.

Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

Schacter, D. L. (2019). Implicit memory, constructive memory, and imagining the future: A career perspective. *Perspectives on Psychological Science*, *14*(2), 256–272.

Schacter, D. L., & Addis, D. R. (2007). On the constructive episodic simulation of past and future events. *Behavioral & Brain Sciences*, *30*(3), 299–351.

Semon, R. (1921). *The Mneme*. London: George Allen & Unwin.
Shoemaker, S. (1970). Persons and their pasts. *American Philosophical Quarterly*, 7, 269–285.

Silva, A. J., Landreth, A., & Bickle, J. (2014). *Engineering the Next Revolution in Neuroscience*. Oxford University Press.

Stich, S. (1978). Beliefs and subdoxastic states. *Philosophy of Science, 45,* 499-518.

Sutton, J. (1998). *Philosophy and Memory Traces: Descartes to Connectionism*. Cambridge University Press.

Von Leyden, W. (1961). *Remembering: A Philosophical Problem*. Duckworth.

Vetere, G., Train, L.M., Moberg, S., Steadman, P.E., Restivo, L., Morrison, F.G., Ressler, K.J., Josselyn, S.A., & Frankland, P.W. (2019). Memory formation in the absence of experience. *Nature Neuroscience, 22,* 933-940.

Werning, M. (2020). Predicting the past from minimal traces. *Review of Philosophy and Psychology, 11,* 301-333.

Westfall, M. (forthcoming). Constructing persons: On the personal-subpersonal distinction. *Philosophical Psychology*.

Wixted, J.T., Mickes, L., & Fisher, R.P. (2018). Rethinking the reliability of eyewitness memory. *Perspectives on Psychological Science, 13,* 324-335.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation.* Oxford University Press.

Yang, W. & Yuste, R. (2017). In vivo imaging of neural activity. *Nature Methods, 14,* 349-359.

Zhang, H., Shi, Y., Jing, P., Zhan, P., Fang, Y., & Wang, F. (2020). Posttraumatic stress disorder symptoms in healthcare workers after the peak of the COVID-19 outbreak: A survey of a large tertiary care hospital in Wuhan. *Psychiatry Research, 294,* 113541.