Forthcoming in the topical collection "Reference and Remembering" in *Synthese*

Engrams as Mental Files

Nikola Andonovski (Centre for Philosophy of Memory, Université Grenoble Alpes)

Abstract:

Engrams—physical memory traces resulting from specific experiences—are the central posits of modern memory science. In this paper, I examine engrams through the lens of the theory of mental files. Integrating evidence from a variety of research programs, I argue that engrams exhibit the core functional properties of mental files. I characterize them as discrete informational structures, formed upon individual experiences of events and causally involved in their subsequent recall. Engrams are plausibly structurally complex in a file-like way, consisting of a stable hippocampal index, which may function as an atomic pointer-like component, and a distributed cortical representation of an event's properties. As such, they afford transmission of content and referential stability during potential content change. Their deployment is constitutive of the capacity for singular reference in episodically remembering particular previously experienced events. This emerging picture of engrams should engender reasonable optimism about the prospects of causal-representational theories of memory.

1. Introduction

Engrams—physical memory traces resulting from specific experiences—are the central posits of modern memory science. Writing about episodic memory, Tulving & Watkins (1975, p. 261) express a sentiment shared by most theorists when they proclaim that an engram, formed upon an experience of an event, "constitutes a necessary condition for the subsequent retrieval of information about the event". Engrams are usually considered necessary for two, closely connected, reasons. Sustaining unique causal chains linking past experiences and present states of recall, engrams play a key role in securing *reference* to the specific events memories are about. Carrying *information* about the events—typically acquired via first-hand experience—engrams afford accurate recall, anchoring memory's reliability. Engrams determine what one is remembering, while making successful remembering possible.[1]

The dual function of engrams is at the heart of Martin & Deutscher's (1966) seminal *causal-representational* theory of memory. On the theory, remembering a previously experienced event requires transmission of content via an engram: a stored representation produced upon the experience of the event, retained in a relatively unaltered form, and causally operative in the production of a current state of recall. As an inner state existing continuously in the interval between a past event and a memory of it, the engram anchors a reference-securing causal chain. The event remembered is the one the experience of which

---

[1] With rare exceptions (e.g., Khalidi 2023), the terms "engram" and "memory trace" have been used synonymously in both the philosophical and the scientific literature. I will follow this practice here. For stylistic reasons, however, I will most commonly use "engram".

occasioned the formation of the currently operative engram. As a representational state, carrying information about the event and making it available for subsequent recall, the engram provides a reliable connection to the past. The event can be accurately remembered because information about it has been previously acquired and safely retained. For Martin & Deutscher, memory is essentially preservative. So, the content of an engram, while fully determining the content of a memory, cannot go beyond the content of the original experience.[2] On this picture, an engram is like a file, created when an event is experienced and stored—in a filing cabinet or the cloud. Upon retrieval, "the full file is pulled from storage and we read off our memories of the experience from it" (Shanton 2011, p. 95).

Intuitively appealing as it may be, this picture has come under a lot of criticism. Pointing to the distributed and superpositional storage of engrams, theorists have become increasingly suspicious of the idea that engrams can sustain unique, and thus reference-securing, causal chains (Sutton 1998; Perrin 2018). Doubts have also emerged about the characterization of engrams as information-bearing or contentful, with a number of alternative—and paradigmatically "minimalist"—conceptions on offer (e.g., Werning 2020; De Brigard 2024a). These have been exacerbated by the increased appreciation of the variety of constructive processes operative at different stages of memory processing. Genuine memories, most theorists agree, emerge only when engrams interact with retrieval cues and other information already available to the subject (Tulving 1983; Moscovitch 2007). More surprisingly, some have even questioned the necessity of engrams, arguing that memory systems can produce an accurate representation of a past event without relying on information acquired upon a particular, direct experience of the event (Michaelian 2016). Whatever one makes of such "postcausal" theories, the emerging consensus seems to be that remembering is not really like reading off memories from a file pulled from storage.

These criticisms are important. Evolving in lockstep with developments in the sciences of memory, they show the limitations of Martin & Deutscher's theory and, arguably, of armchair theorizing in general (Zemach 1983). Yet, their presentation tends to obscure the resilience of the core idea behind the causal-representational approach. At the heart of the idea is the functional duality of engrams as reference-securing and information-bearing physical states connecting the present to the past. The recent revival of the "search for the engram" has revitalized the commitment to discrete engrams with distinguishable, and possibly empirically tractable, causal histories (Josselyn & Tonegawa 2020; Guskjolen & Cembrowski 2023). In this literature, engrams are routinely characterized as carrying information about, or representing, states of the environment, thus affording reliable recall. These developments suggest that, despite the shortcomings

---

[2] Here I follow the characterization in Michaelian (2011, 2016). As far as I can see, Martin & Deutscher (1966) believe that *any* state of remembering, regardless of accuracy, must satisfy this "content-matching" condition. What kind of modality is at play in this claim is a difficult question, which I briefly address in Andonovski (2021a), but which certainly deserves a more thorough treatment.

of Martin & Deutscher's theory, the core idea behind it may well be worth saving. Examining its status—and indeed: promise—requires not only engagement with the rapidly developing science of memory but also an examination of the conditions for positing engrams with the requisite functional profiles. Such work may be the basis for new, empirically adequate, causal-representational theories of memory.

In this paper, I examine engrams through the lens of the theory of mental files, developed originally in the philosophical literature on reference. Integrating evidence from a number of research programs, I argue that engrams exhibit the core functional properties of mental files, guiding mnemic reference to previously experienced events and storing information about them. The emerging picture of engrams as file-like informational structures, causally involved in remembering the past, should engender reasonable optimism about the prospects of causal-representational theories. The paper is structured as follows. In section 2, I introduce the notion of a mental file, characterizing mental files theories as working empirical hypotheses. In section 3, I present the main hypothesis of the paper and provide evidence in its support, examining the formation, informational structure, and causal relevance of engrams. In section 4, I take stock, positioning the files view in the literature and outlining the major challenges to developing a naturalist causal-representational theory of memory. Section 5 is the conclusion.

2. Mental Files

2.1. Mental Files: Singular Reference and Structural Complexity

The notion of a mental file entered the philosophical toolkit in the debates about the nature of *singular thought.* A singular thought is a thought that is, in some intuitive sense, directly about an individual. The paradigmatic form of such thought is occasioned upon a perceptual encounter with an object or person: my thought that *that* woman (I can point to) is wearing a green blazer seems to be directly about the woman in question. Singular thoughts are typically contrasted with *general* thoughts about properties or kinds—e.g., my thought that green blazers are fashionable—but also with thoughts which refer to individuals but only as objects of thought that satisfy certain descriptions. My thought that the writer of *White Teeth* loves green blazers is about an individual, yet seemingly only in an attenuated sense. While admitting the intuitive appeal of this distinction, *descriptivists* ultimately reject it, arguing that mental reference to individuals requires that they are represented as satisfying some descriptive conditions (Searle 1983; Nelson 2002). *Non-descriptivists*, in contrast, argue that we can have thoughts about individuals solely in virtue of standing in some direct relation to them—a relation paradigmatically characterized as one of *acquaintance* (Evans 1982; Recanati 2012).

Mental files theories of singular thought appeal to the deployment of a particular kind of representation: a mental file (Bach 1987; Jeshion 2010; Crane 2011). On the non-descriptivist theory of interest to us here, singular thinking constitutively involves the tokening of such a representation: to entertain a singular thought about an individual *is* to deploy a file about it (Recanati 2012, 2016).[3] Mental files, accordingly, are mental representations whose functions are to guide reference to individual entities and to collect, store, and retain information about them. As Recanati (2012, p. 35) explains:

> Mental files are "about objects": like singular terms in the language, they refer, or are supposed to refer. They are, indeed, the mental counterparts of singular terms. What they refer to is not determined by properties which the subject takes the referent to have (i.e., by information—or misinformation—in the file), but through the relations on which the files are based. The referen[t] is the entity we are acquainted with (in the appropriate way), not the entity which best 'fits' [the] information in the file.

Mental files are thus cognitive and representational structures whose deployment makes it possible to entertain singular thoughts about entities *as* individual entities and not simply as possessors of properties that satisfy certain descriptions.

A mental file is structurally complex, consisting of the *file* itself—a cognitive particular formed upon some encounter with an entity and persisting across time—and the *body of information* stored in it—i.e., the contents of the file, representing attributes predicated of its referent. Crucially, a file can retain its identity even if its contents change.[4] This structural complexity is of key importance, anchoring a file's functional role. Building on the characterization in Recanati (2012) and Murez et al. (2020), we can identify three central characteristics of structurally complex files:

(1) The *reference* of a file is determined relationally—via relevant causal-historical or informational relations to individual entities—rather than via descriptions. Recanati (2012) speaks of *epistemically rewarding* (ER) relations, enabling the subject to gain information from the entities they stand in relations to.

(2) Informational entries are associated with certain files in virtue of *common ER relations.* Specifically, information gets filed together because it is acquired via the same ER information channel.

(3) Files are not, or at least need not be, *accessed* via their informational contents. Rather, a file may be activated directly.

A few words about each of the characteristics. First, while perceptual acquaintance is the paradigmatic ER relation, Recanati (2012) posits a variety of ER relations—perceptual, mnemic, testimonial etc.—

---

[3] In the rest of the paper, I will rely exclusively on Recanati's (2012, 2016) influential "indexical" mental files theory. It is nevertheless worth noting that the notion of a mental file can be, and indeed has been, appealed to by theories with different commitments and explanatory priorities.

[4] Unless specified otherwise, any reference to files in the rest of the paper will concern mental files specifically.

corresponding to different types of files, each of which anchored on a causal chain permitting a flow of information between the thinking subject and the relevant entity. Given this liberal approach, it is an open question whether a *unified*, and suitably restricted, account of ER relations is forthcoming—a concern exacerbated by the apparent cognitive richness of many of the posited relations (Hansen & Rey 2016). Second, informational entries are filed together when acquired via the same ER relation—e.g., information about the size and shape of an object will be co-filed upon a perceptual encounter with it. Mental files thus serve as repositories of information gained in a specific way.[5] Third, as structurally complex particulars, files can be accessed directly—i.e., not via its informational entries. What such direct access involves, *psychologically,* is an open question, with various access procedures likely to (co)exist in different cognitive domains.

In virtue of these characteristics, mental files are able to function as *non-descriptive modes of presentation*—or senses—of given referents. One may have multiple files for the same referent—based on different ER relations—and may thus rationally attribute contradictory properties to the same referent. Hence, I may believe that the woman I am pointing to is wearing a green blazer, while believing on the basis of testimony that the author of *White Teeth* never wears green, simply in virtue of tokening two different files for what turns out to be the same individual. Indeed, "if there are two distinct files, one associated with [the woman I am pointing to] and the other with [the author of *White Teeth],* then there are two distinct senses, *even if the information in the two files is the same*" (Recanati 2012, p. 41). Relatedly, a subject can successfully refer to an entity—by deploying a file with the proper causal history—even if they mistakenly attribute to it properties that the referent entity does not possess. Combining the virtues of descriptivist and non-descriptivist approaches, mental files thus aim to account not only for singular reference and information coordination but also for a thought's cognitive significance.

2.2. Mental Files Theories as Empirical Hypotheses

Mental files theories appeal to representational structures with characteristic functional profiles to account for the phenomena of interest. This explanatory strategy has generated some concern, with theorists suspecting that the appeal—typically a result of *a priori,* armchair theorizing—provides only a metaphorical way of characterizing the cognitive clustering and coordination of information (see, e.g., Goodman & Gray 2022). If mental files theories are to offer more than a metaphor, then mental files must be *psychologically real* cognitive particulars, posited not only on theoretical but also on empirical grounds. As Murez et al. (2020, p. 110) explain:

---

[5] It is worth noting, however, that Recanati (2012) also posits more complex files, such as so-called "encyclopedia entries", which abstract from *specific* ER relations, yet are still about individual objects.

> If files are psychologically real, then merely sketching an account in which they fit the task-description of essentially singular representations a priori is insufficient... The issue is not how we use the technical expression "mental file", but whether an empirically well-motivated notion in its vicinity can explain or support, rather than merely *label*, the distinction between singular mental representations and general or descriptive ones.

Mental files theories are thus best understood as *empirical hypotheses* about the existence and functional properties of mental files as well as the role they play in information coordination and singular thought. For our purposes, we can distinguish between *global* and *local* mental files hypotheses. On a global hypothesis, mental files form a psychological natural kind, whose deployment is constitutive of the (human) capacity for singular thought in *all* of its paradigmatic forms. While arguably a better fit for the explanatory ambitions of most mental files theorists, such a global hypothesis—as Murez et al. (2020) document in their comprehensive treatment of the issue—seems poorly supported by the available evidence. Given that the category of singular thought is typically delineated in folk-psychological, intuitive terms, this is probably not very surprising (Hansen & Rey 2016). This verdict, however, should not militate against the consideration of *local* hypotheses, not committed to the existence of mental files as a unified psychological kind. A local hypothesis identifies a class of mental representations, exhibiting the characteristic functional profile of mental files, proposing that their deployment is constitutive of the capacity for singular reference in some, theoretically delineated, cognitive activity—e.g., person or object representation (Murez & Smortchkova 2014; Siegel 2022).

The aim of a local hypothesis is to reveal robust, and explanatorily significant, functional similarities between files and a candidate class of mental representations. While the selection of such class will typically be driven by specific empirical developments, the hypothesis should ideally be supported by evidence from a variety of scientific disciplines. This stance exemplifies a commitment to a form of methodological naturalism, on which theorists "take their lead" from productive research programs while aiming to develop a general picture of a phenomenon of interest (Maddy 2007; Andonovski & Michaelian, forthcoming). Such a picture—in our case: of singular reference in a given cognitive domain—should integrate scientific insights and bring them into contact with philosophical notions and concerns, yet remain subject to further amendment and revision. A local hypothesis, nevertheless, should *not* be expected to provide a "smooth" reduction of mental files or to accommodate features typically characterized in an abstract of metaphorical way.[6]

---

[6] A reader may, and a reviewer does, worry that singular reference is not a naturalistic notion. My response to this worry is twofold. First, the version of naturalism endorsed here is of a *methodological* kind, and it is characterized by features of philosophical practice such as engagement with empirical inquiry, privileging of scientific evidence, and a general suspicion of armchair theorizing (Andonovski & Michaelian, forthcoming). Importantly, the commitment to this kind of naturalism does not issue *a priori* constraints on, or guidelines about, which notions are naturalistic (unlike, e.g., some forms of neo-Quinean metaphysical naturalism). Second, in this explanatory project, singular reference is

The rest of the paper is concerned with one such local hypothesis, positing the deployment of engrams as constitutive of the capacity for singular reference in episodic memory. As we will see, "translating" the commitments of mental files theories to the idiom of memory science does indeed constitute the principal challenge.

3. Engrams as Mental Files

After introducing engrams in 3.1., I use 3.2. to present the main hypothesis of the paper. In the remainder of section 3, I provide evidence in its support. In 3.3., I examine engram formation, looking at core properties of event perception and encoding. In 3.4., I provide evidence for the causal relevance of engrams in episodic recall. In 3.5., I characterize them as structurally complex in a file-like way and thus capable of sustaining referential stability.

3.1. Engrams: A functional profile

Engrams are neural memory traces, the physical substrates of stored information resulting from past experiences and affording subsequent memory expression. Characterized as the vehicles of memory retention, engrams have been a prominent feature of thinking about memory since at least Plato—theorized about under different names and with diverse conceptual tools and metaphors (Draaisma 2000). Coining the term "engram", Semon (1921) made explicit the commitment to a view of the engram as a lasting *neural* change in response to a learning experience, thus inaugurating the modern neuroscientific search for it.[7] After a decades-long hiatus—caused by a variety of theoretical, methodological, and ideological obstacles—the search was recently revived, spurred by technological advances and the development of novel tools for probing and manipulating brain function at neural, synaptic, and molecular levels (Josselyn et al. 2015; Tonegawa et al. 2015). Key among these was the development of optogenetics, which allowed researchers to track and control the activity of individual neurons with light (Deisseroth 2010).

While characterizations of the engram vary, there is relative agreement about its broad functional profile. Four key properties, identified by Josselyn et al. (2015), are worth highlighting:

---

the explanandum, not the explanans. So, while it may in fact turn out that singular reference, strictly speaking, does not exist, whether it does is an open empirical question, to be settled by examining evidence from a variety of disciplines. The reviewer appeals to Openshaw (2023), who may be read as arguing that psychofunctionalist theories should not be seen as aiming to account for "referential remembering". This is *not* my preferred interpretation of the article, but exegetical issues aside, the endorsement of an anti-referentialist psychofunctionalism does not entail that reference is not a naturalistic notion, in any interesting sense. As Openshaw (2023, p. 297) readily admits, any complete theory of remembering will have to say *something* about "the reference question".

[7] For the details of Semon's complex theory of memory and the role engrams play in it, see Schacter et al. (1978).

(1) *Experience-dependence:* an engram is a lasting neural change resulting from a specific experience or event.

(2) *Information storage:* an engram carries information, or content, about the relevant experience or event.

(3) *Dormancy:* an engram may exist in a dormant state between the processes of encoding and retrieval; as such, it is independent of them.

(4) *Behavioral expression:* an engram may be behaviorally expressed through interaction with various retrieval cues. It thus has the potential for "ecphory"; i.e., memory retrieval.

As Robins (2023) points out, this functional characterization is skeletal, less a synthesis of existing knowledge than a supportive platform for budding empirical research—conscientiously constructed but provisional. With this research in its early years, and with the extent of theoretical disagreement about the nature of engrams, such a characterization is likely necessary. Memory scientists disagree not only about the precise mechanisms of memory storage (Poo et al. 2016; Gershman 2023) but also about the organization, diachronic stability, and long-term location of engrams (Winocur & Moscovitch 2011; Barry & Maguire 2019). Despite this, we can discern in this functional profile some important characteristics of engrams, worth a brief discussion.

Engrams are the physical substrates of retained information. A commitment to this claim involves two, conceptually distinct, components (Robins 2023). Engrams are *vehicles*: physical (e.g., neural) structures that have causal powers and play important roles in the generation of recall and memory-based behavior. They are also *information-* or *content-* bearing: they support the retention of information, acquired in past experience(s) and subsequently employed in remembering. The belief that stored information is required to account for key phenomena—such as, e.g., the artificial "retrieval" of otherwise inaccessible memories (e.g., Guskjolen et al., 2018)—is widespread in the memory sciences, despite some recent dissenting views (e.g., Brette 2019). As (2) illustrates, indeed, engrams are widely considered to be *representational* structures, carrying information about a past event or learning experience.

As vehicles of memory retention, engrams are taken to be *discrete* representational structures. What such discreteness involves has been a matter of some controversy and confusion. Following Dietrich & Markman (2003) and Maley (2011), we can think of discreteness as a property of representational schemes, employed by representing systems. A representational scheme is discrete iff it contains multiple representations, each of which is distinct from other representations in the scheme (i.e., uniquely identifiable), and there are gaps between the possible representations.[8] With some terminological liberty,

---

[8] A discrete representational scheme, i.e., is one that is not *continuous*. It is worth noting, however, that discrete representational schemes can nevertheless be *dense* (see Maley 2011, p. 125).

we can characterize an individual representation as discrete if it is a part of a discrete representational scheme (e.g., letters of the alphabet, rational numbers). Discrete representational schemes, it has been argued, allow cognitive systems to discriminate distinct conditions in their inputs and categorize those inputs accordingly (Dietrich & Markman 2003). Engrams are thus taken to be discrete representations, in this relatively minimal sense, insofar as they are hypothesized to be parts of discrete representational schemes employed by memory systems. Each engram is uniquely identifiable and distinct from other representations; indeed, engrams are typically taken to sustain unique causal chains linking states of recall to past experiences. As a result, engram-mongering memory systems can discriminate between, and represent, distinct events.[9] While engrams are considered discrete vehicles of memory retention, however, they are not typically taken to *be* memories. Rather, as specified in point (4) above, engrams are behaviorally expressed through interaction with various retrieval cues. Only upon such interaction do genuine, full-fledged memories emerge (cf. Moscovitch 2007). Why engrams are insufficient for the emergence of memories and how exactly we should characterize the causal and informational role they play in such emergence are difficult open questions, to be examined on a different occasion.

The number of open questions does not make engram theories empty or trivially true. The commitment to discrete neural vehicles carrying information about specific events, is a substantive one. Indeed, much of the recent dissatisfaction with causal-representational theories, in the philosophical literature at least, has been due to skepticism about the existence of discrete neural vehicles supporting memory retention. This is precisely what makes the budding scientific study of retention mechanisms important. Nevertheless, as we formulate our local hypothesis and follow the evidence in an attempt to justify or at least motivate it, we should do so with appropriate modesty. Not only is engram theorizing, as indeed most neuroscientific research, fraught with fallibility, but the sheer variety of competing hypotheses necessitates that we keep both feet firmly on the ground.

3.2. Engrams as Mental Files: The Hypothesis

When theorists like Tulving & Watkins (1975) characterize engrams as necessary for memory retrieval, they have in mind *episodic memory:* the capacity to remember events or experiences from one's personal past, such as a past birthday party or a walk in the woods. The focus on episodic memory, similarly exemplified in philosophical work, is not surprising. Engrams are characterized as neural structures resulting from, and carrying information about, *specific* experiences organisms undergo. While they may

---

[9] We should distinguish discreteness from *locality.* A representational scheme is localist iff each entity is represented by activity in a single computing element. On a prominent conception, representations are *distributed* iff they are not local (Van Gelder 1991). Importantly, engrams that are distributed (in *this* sense) can nevertheless be discrete—i.e., distinct and uniquely identifiable by the system. I return to a richer notion of distribution, linked to the idea of superposition, in 3.5.

contribute to other forms of memory—e.g., remembering African capitals or how to shoot free throws—they are, in the eyes of most theorists, causally and explanatorily central to remembering particular past events or experiences.[10]Accordingly, my focus on this paper is exclusively on episodic memory as a theoretically delineated cognitive kind. I remain neutral on whether engrams are required for other kinds of memory retention or retrieval, an issue that has caused some controversy in the literature (Thompson 2005; Eichenbaum 2016).

Episodic memory, then, is the capacity to remember events from the personal past. In his groundbreaking proposal, Tulving (1983) argued that this capacity is underlaid by a specialized memory system. When functioning properly, the system—tentatively defined as a set of dedicated neurocognitive processes operating in correlation—was taken to afford first-hand knowledge of previously experienced events.[11] The nature of the recollective experience was seen as reflecting this, with subjects experiencing remembered events *as* previously experienced. Tulving (1985) labeled the kind of consciousness conferring this feeling "autonoetic"; i.e., self-knowing.[12] Despite a number of interim developments, the system-centric view and the centrality of the notion of autonoetic consciousness—as well as the closely associated *mental time travel*—have remained core features of contemporary episodic memory science (Addis 2020; Ranganath 2022).[13] This preliminary characterization of episodic memory allows us to introduce our local hypothesis:

> *Engrams as Mental Files (EMF).* The deployment of engrams, a class of mental representations that exhibit the characteristic functional properties of mental files, is constitutive of the capacity for singular reference in episodically remembering particular previously experienced events.

On the hypothesis, engrams are psychologically real cognitive particulars with file-like properties. Their deployment in episodic remembering secures singular reference to previously experienced events. Engrams store information about the properties of such events. The information is typically acquired via a direct experience of a relevant event and subsequently utilized in episodically remembering it. The tokening of engrams accounts for an episodic memory's unique cognitive significance.

---

[10] On the standard model, these are instances of semantic and procedural memory, respectively (Squire 2004).

[11] While episodic memory was hypothesized to be dissociable from semantic and procedural memory, the systems were nevertheless taken to be functionally interdependent and frequently interacting (see Tulving 1983, Andonovski 2023).

[12] For Tulving (1983), this was one of the reasons episodic memory was *uniquely* human, a conclusion not shared by the majority of memory theorists.

[13] For a recent dissenting view, see De Brigard (2024b), who argues that the construct of autonoetic consciousness lacks validity and that, consequently, episodic memory should be characterized independently of it.

Let's unpack these claims a bit. In remembering episodically, subjects entertain singular thoughts about past, previously experienced events. An episodic memory, emerging through the interaction of an engram with available retrieval cues, is—in the theoretically relevant sense—*directly* about an individual event; not simply as possessor of certain attributable properties. This capacity for episodic singular reference is the primary explanatory target of the EMF hypothesis. On it, what accounts for such reference is the tokening of a file-like engram with a unique causal history. Episodic singular thought *is* engram-supported thought.[14] Engrams sustain epistemically rewarding relations to individual events, permitting information flow from past experiences to present memories of them. These ER relations are characteristically episodic in that they are linked to the normal functioning of the episodic memory system and/or the constituent (episodic) processes. A prototypical episodic ER relation will involve first-hand perceptual experience of an event. Yet, as we will see, there may be different kinds of episodic ER relations. Some heterogeneity in episodic causal-informational chains is thus to be expected.

Informational entries are associated with particular engram-files in virtue of common episodic ER relations. Specifically, information about different properties of an event gets filed together—or "bound" in an integrated representation—because it is acquired via the same ER information channel. In this way, an engram functions as a repository of information about an event gained in a specific, canonically episodic, way. It is subsequently behaviorally expressed through interactions with internal or external retrieval cues. Yet, as a file-like structure, an engram need not be—and perhaps is typically not—accessed via its informational entries. The deployment of an engram in recall constitutes a specific mode in which an event is presented in memory—paradigmatically: as previously experienced—with specific behavioral and phenomenological signatures. Importantly, a subject can successfully refer to an event in memory—by deploying an engram-file with the appropriate causal history—even if they mistakenly attribute to it properties it does not possess, relying on false information contained in the engram. Such misattribution may occur as a result of misperception, information loss, and perhaps even information updating. The account also leaves open the possibility that the subject may have multiple mental files for the same referent event: e.g., a testimonial and/or semantic memory file in addition to the engram-file. Only when the engram-file is tokened does one have a characteristically episodic singular thought about the event.

In the following three sections, I provide empirical support for EMF, characterizing engrams' formation, structure, and the role they play in episodic recall.

---

[14] Singular reference in episodic memory is the primary explanatory target of EMF, hypothesized to be *constituted* by the tokening of engrams with file-like properties (structural complexity, discreteness, contentfulness etc.). Engrams are necessary for episodic remembering in that they make such singular reference possible. The resultant picture, I suggest in 4.2, may form the basis of a naturalist causal-representational theory of memory, to be developed on a future occasion. I am grateful to an anonymous reviewer for prompting me to add this clarification.

3.3. Event Perception, Engram Formation, and Episodic ER Relations

Events are the prototypical referents of episodic memories. On the proposed account, the reference of a memory is determined relationally, with engrams sustaining ER causal relations, permitting information flow from a past experience to a current state of recall. The remembered event is the one the experience of which occasioned the formation of the engram currently tokened and behaviorally expressed in a memory. Here, I examine engram formation, looking at core properties of event perception and encoding and their downstream effects on memory recall. I characterize such formation as automatic, plausibly inferential, and involving the binding of information about such properties in integrated event representations.

Events correspond to natural and functionally important units of cognition. In the absence of an agreed-upon definition of "event", theorists have adopted a variety of approaches and operationalization procedures, often characterizing events simply *as* the happenings that correspond to meaningful units of cognition or experience (see Yates 2023).[15] For our purposes, we can tentatively adopt the characterization offered by Zacks & Tversky (2001), according to which an event is a segment of time at a given location taken by the subject to have a beginning and an end. Three components of it are worth highlighting. First, events have various properties: spatial and temporal but also objectual, structural, causal, agential etc. Second, events are considered to have relatively determinate beginnings and ends, even if the boundaries between them are porous and, to an extent, manipulable. Third, events are composed by, and can be decomposed into, "smaller" (sub-)events, each of which *may* correspond to a meaningful unit of cognition. (E.g., walking to work may be composed of leaving the house, stopping by at the coffee shop, meeting a friend, reaching the office building etc.)

The prototypical episodic ER relation is anchored in perceptual experience. Half a century after the pioneering work of Newtson (1973), the evidence for segmentation of perceptual experience into discrete events, at multiple scales, has accumulated (for reviews, see Zacks 2020; Wang et al. 2023). Studies show that, despite some individual differences, people tend to segment their perceptual experiences in very similar ways, agreeing where the "natural" boundaries between events are. Yet, they appear to do this in a flexible manner, capable of responding to task demands and manipulating the grain of segmentation (e.g., Bailey et al. 2017). Indeed, when asked to segment at multiple scales, people seem to organize event representations hierarchically, with fine-grained events grouped into coarser grained ones (Hard et al. 2006; Zacks et al. 2007). Nevertheless, there is now considerable evidence that at least some forms of perceptual

---

[15] In the event cognition literature, "event" is sometimes used to stand for the internal, functional unit of cognition (e.g., an event representation) rather than for the external occurrence such a unit corresponds to. In this paper, I will use "event" exclusively for the external occurrence. See the discussion at the end of this section, however.

event segmentation are automatic and independent of cognitive load (Kurby & Zacks 2008). Studies have implicated a number of brain regions in segmentation, with lower-level perceptual regions shown to be responsive to fine-grained events, while higher-level multimodal regions and the hippocampus appear responsive to more coarse-grained, longer events (Baldassano et al. 2017; Geerligs et al. 2022).

Crucially, perceptual event segmentation has downstream effects on recall, supporting the idea that event representations are functionally important units in episodic memory. A variety of studies have reported superior recall for information presented at event boundaries (Boltz 1992; Swallow et al. 2009), with source memory for boundary information similarly shown to be superior (Heusser et al. 2018). Moreover, and perhaps more importantly, segmentation affects sequential binding in long-term memory. In a landmark study, Ezzyat & Davachi (2011) showed superior recall for information within an event compared to information across event boundaries, while Dubrow & Davachi (2013) reported a similar result for temporal order memory. Event segmentation also affects prospective and retrospective duration judgments, with temporal intervals containing more event boundaries consistently estimated to be longer (Faber & Gennari 2015). In free recall, the presence of boundaries between items also modulates the likelihood of serial transitions between them (DuBrow & Davachi 2016; Heusser et al. 2018).

By examining event segmentation, we have zeroed in on the beginning of the hypothesized causal chain linking perceptual experience and subsequent recall. On the standard models, active event representations resulting from, and guiding subsequent, segmentation are held in working memory—as, e.g., in Radvansky & Zacks' (2017) influential Event Horizon Model.[16] Since Atkinson & Shiffrin (1968), it has been commonplace to posit that such representations are subsequently transferred to a long-term memory store—and appropriately transformed in the process—with the transfer occasionally thought to coincide with encoding or the formation of the engram, strictly speaking (e.g., Tulving 2002).[17] Yet, what matters for us, at this point, is not the precise relation between perceptual and memory processing or the existence of a distinct working memory store. It is rather the existence of uninterrupted causal-informational chains from perceptual experiences of events to subsequent states of recall, which can anchor epistemically rewarding relations and *are,* at the least in the long run, sustained by engrams. Accordingly, my focus here is on the way such a causal relation to an event, permitting the acquisition of information and storage in memory, is established in perceptual experience.

---

[16] A reviewer asks whether these event representations are thus distinct from object files, as characterized by (e.g.,) Green & Quilty-Dunn (2021). The two kinds of files indeed share important functional properties, such as structuring automatic working memory processes (see Murez et al. 2020). Despite this, they differ in their referential domain (events vs objects) and—if the arguments below are on the right track—long-term maintenance conditions. Further work is needed to examine the similarities and differences between the two kinds of files in more detail.

[17] Some theorists have steadfastly rejected this idea, insisting that encoding is essentially a perceptual process (e.g., Craik 2020).

Event segmentation appears to depend on the detection of changes in a variety of features. In the typically highlighted cases, these are external, and characteristically low-level: e.g., spatiotemporal location, motion, color, and sound (for a review, see Radvansky & Zacks 2014). It may thus be tempting to characterize event segmentation as a brutely causal process in which the simple detection of low-level changes in stimulus features triggers the opening of an engram-file about an event—a proposal along the lines of Pylyshyn (2007). We have to proceed with care, however. The evidence that changes in internal—e.g., in affective, motivational or goal—states play an important role in event perception and memory has steadily accumulated (for a review, see Wang et al. 2023). Indeed, a recent study has shown that internally generated changes in goal states—plausibly guided by the *descriptions* of varying task demands—can create event boundaries even in the absence of any change in the external stimulus (Wang & Egner 2022). Accordingly, leading theories have characterized the process of segmentation as inferential and as likely involving prior representations of an event's properties—e.g., in multidimensional event models (Zacks 2020; Wang et al. 2023). This does not undermine EMF. The hypothesis requires the existence of an uninterrupted causal chain, permitting information flow from an experienced event to a subsequent state of recall. It does *not* require that the individuation of an event in perceptual experience, triggering engram formation, is brutely causal or guided only by non-representational detection mechanisms. Indeed, outside of early vision and a few other input systems, reference-fixing mechanisms in human thought are unlikely to resemble Pylyshyn's FINSTs. Any characterization of the notion of acquaintance in the context of mental file theories, however strong its debt to Russell, should reflect this (Recanati 2012; Murez et al. 2020).[18]

Crucially, informational entries about properties are filed together in virtue of being acquired via the same causal-informational channel to an event thus individuated. Attention appears to be involved in the early stages of the process, being automatically allocated to features taken to belong to the same event and reset at event boundaries (De Freitas et al. 2014; Yousif & Scholl 2019). Downstream, co-filing is supported by the construction of integrated engram-files, with the hippocampal formation—employing a dynamic neural code to bind various informational entries—believed to play a key role (Sugar & Moser 2019; Ross & Easton 2022). We saw above that segmentation affects sequential binding in long-term memory. Further behavioral evidence for co-filing is provided by recent studies showing that event information in episodic memory tends to be retrieved and forgotten in an all-or-none manner, with the binding of features at encoding having lasting effects on subsequent recall (Horner et al. 2015; Joensen et al. 2020). Extending this work, Mahr et al. (2021) show that information about the location and time of day—but *not* information about temporal orientation—of an event is tightly integrated in episodic recall.

---

[18] Indeed, some theorists feel that, in light of this change, the notion of acquaintance has ceased to be illuminating (e.g., Hansen & Rey 2016). This may, nevertheless, end up being a terminological dispute.

These results suggest that the association of informational entries with specific events is stable over time, supporting relatively independent ER causal chains from distinct events to acts of recall.

Before we move on, a few words about the possible heterogeneity of episodic ER relations. On the picture presented thus far, it would be easy to restrict the referents of episodic memory to external, characteristically short, events. Yet, episodic memories can seemingly be about longer, more complex events (e.g., concerts, job interviews, family vacations) and may incorporate various internal elements (e.g., affective or motivational states). It is, of course, an open empirical question whether, and to what extent, these memories are supported by the same underlying mechanisms. Nevertheless, recent work suggests that the account has the resources to accommodate them. Perceptual experience, as we have seen, is typically segmented at multiple temporal scales, with event representations organized in a hierarchical way. The evidence shows that all-or-none retrieval and forgetting effects occur across this hierarchy but are underlaid by different binding mechanisms (Collin et al. 2015; Andermane et al. 2021). Indeed, some forms of binding are likely decoupled from immediate perception, with information acquired in perceptual experience but only subsequently bound in narratively structured event representations (Cohn-Sheehy et al. 2021). Internal elements of experiences can be incorporated at multiple scales. Recent evidence suggests not only that internal states affect segmentation but also that hippocampal mechanisms underlie the incorporation of affective, motivational, and self-related information in memory (McKenzie et al. 2014; Zeithamova et al. 2018; Ross & Easton 2022). Indeed, episodically remembered events may, in principle, consist entirely of internal elements.[19] These considerations leave open a number of difficult questions.[20] Yet, assuming the reliable individuation of events and the relative stability of informational channels to them, the heterogeneity of ER relations does not present a principled obstacle to EMF.

3.4. *Causal Dependence, Discreteness, and Dynamicity*

The existence of an episodic ER relation requires the causal and informational dependence of successful recall on the prior encoding, storage, and maintenance of the relevant content. In the previous section, the focus was primarily on the informational aspect of the dependence: we saw how the binding of event features at encoding affects the organization of information in long-term memory, with lasting effects on subsequent recall. In this section, I examine the causal aspect more directly, focusing on the role engram-files *qua* physical structures play in the generation of recall. I look at evidence from the subsequent memory

---

[19] This will be reflected in the memories' correctness conditions. Some philosophers (e.g., Bernecker 2010) indeed believe that there are two distinct dimensions of evaluation of episodic memories: *truth* (correspondence to the external event) and *authenticity* (faithfulness to the past internal experience). I bracket this issue here.

[20] Not the least of which concerns the nature of the causal-information link between a complex event representation and the (multiple) event(s) it corresponds to (e.g., a family vacation and its "sub-events"). Making progress on issues like this may require knowing much more about how the relevant cognitive systems work.

effect paradigm and the exciting new research on engram capture and manipulation. The latter, in particular, is of vital importance for establishing engrams as discrete information-bearing structures. Such discreteness, nevertheless, does not entail that an engram's informational content, or even structural basis, cannot change over time.

On EMF, episodic recall requires an uninterrupted causal chain, permitting information flow from an experienced event to a subsequent state of recall. Such a chain will consist of a set of dependence relations—counterfactual or probabilistic—with particular states/processes dependent on earlier states/processes in the chain. The endpoint of the chain—episodic recall—is hypothesized to be dependent on a variety of earlier states involved in encoding, consolidation, and storage.

Much of the evidence for the causal dependence of recall on *encoding* comes from the systematic investigation of so-called "subsequent memory effects" (SMEs). In the paradigm, researchers examine the variable neural activity at an encoding phase of an experiment as a function of subsequent retrieval success. Studies have identified reliable differences in neural signals—i.e., SMEs—relative to subsequently remembered and forgotten items. Particular SMEs have been found in a number of brain regions and in a variety of characteristically episodic tasks, and have been linked to specific encoding mechanisms such as attention and associative binding (Jenkins & Ranganath 2010; Xue et al. 2010; Aly & Turk-Browne 2016). Importantly, if recall is causally dependent on a particular encoding activity, then the manipulation of the activity should affect memory performance, *ceteris paribus.* With direct manipulation of encoding activity in humans difficult and ethically problematic, researchers have resorted to assessing its causal relevance by statistical analysis from large-scale observational data. This is the approach taken by Weidemann & Kahana (2021), who examined SMEs in a word list free recall task, while statistically controlling for the effects of a comprehensive set of external factors. They showed that SMEs reflected endogenous neural activity predictive of recall success. Rubinstein et al. (2023) reported a similar result, with specific high-frequency activity in the hippocampus correlating strongly with item recall. These results provide preliminary evidence for the causal dependence of successful recall on prior encoding.[21] Yet, important concerns remain. SME studies do not involve direct manipulation of the neural activity hypothesized to be causally relevant, leaving room for the possible influence of unaccounted-for external factors. They are also limited to establishing the dependence between encoding and retrieval, providing no insight on the nature of the, presumably engram-sustained, causal chain connecting them.

These concerns are partially assuaged by the reinvigorated research program targeting engram capture and manipulation. The organizing idea behind this program is that engram formation involves the

---

[21] See, however, Halpern et al. (2023), who were unable to identify any neural signals that reliably predicted successful recall, after statistically controlling for relevant external factors.

modification of the strength of synaptic connections between populations of neurons—neural *ensembles*—active during encoding (Hebb 1949; Poo et al. 2016). Such modification is hypothesized to make subsequent retrieval possible by increasing the likelihood that relevant patterns of activity in the ensembles will be recreated upon interaction with retrieval cues. Researchers employ a variety of novel molecular and transgenic methods—sometimes grouped under the term "engram technology" (Tonegawa et al. 2015)—to capture, tag, and manipulate neural ensembles active at the time of learning. By doing so, they aim to provide direct evidence for the causal import of such ensembles—conjectured to house individual engrams or engram components—on memory performance. Engram-technological manipulations are most commonly employed in rodent conditioning studies and usually target hippocampal and amygdalar populations.

Two families of studies are worth highlighting here (for reviews, see Josselyn et al. 2015; Tonegawa et al. 2015; Guskjolen & Cembrowski 2023). Loss-of-function studies involve the selective inhibition of neurons active during encoding with the aim of preventing their reactivation and resultant memory expression. Engram-technological advances, such as gene labelling and optogenetics, have allowed researchers to do this in a localized manner, tagging and selectively silencing engram neurons. Denny et al. (2014), for example, tagged populations of neurons in the dentate gyrus and CA3 region of the hippocampus active during encoding of a fear-inducing context in mice. The optogenetic silencing of the tagged neurons prevented the expression of the relevant contextual-fear memory, with performance deficits observed weeks after training. Loss of function, upon silencing or ablation of engram neurons in different regions, has been reported in a growing number of studies (e.g., Rashid et al. 2016; Zhou et al. 2019). Gain-of-function studies are arguably more attention-worthy. They typically involve artificial reactivation of a neural ensemble with the goal of triggering the behavioral expression of a memory. In a landmark study, Liu et al. (2012) tagged dentate gyrus neurons active during auditory fear conditioning in mice. Remarkably, when these neurons were optogenetically activated in a novel context, previously unpaired with a foot shock, the mice exhibited fear, manifested in freezing behavior. The approach has also been used to reactivate memories for rewarding events (Redondo et al. 2014; Ramirez et al. 2015). In a particularly intriguing demonstration that optogenetic stimulation of tagged neurons can induce memory expression, Guskjolen et al. (2018) recovered seemingly "lost" infant memories in mice up to 3 months following training.

Engram capture and manipulation studies establish the causal import of neural ensembles associated with particular engram-files. In loss-of-function studies, the targeted inhibition of an ensemble leads to a decrease in, or indeed complete elimination of, the behavioral expression of a memory. Successful recall—believed to correspond to such expression—thus appears to be causally dependent on the activity of an engram ensemble. Moreover, gain-of-function studies appear to demonstrate that the artificial activation of an ensemble is *sufficient* for the behavioral expression of a specific memory (Josselyn et al. 2015). While this finding may be *prima facie* surprising—engrams, recall, have been traditionally considered *in*sufficient—we should observe that, by bypassing natural retrieval processes, optogenetic activation may effectively subsume engram-cue interaction.[22] It is worth noting in this context that the target activity of a neural ensemble need not be taken to constitute the complete cause of a mnemic behavior; it only has to be a *part* of the causal nexus producing it (cf. Cao 2022).

As Najenson (2021) and Robins (2023) emphasize, however, *specificity* is key. Engrams are hypothesized to be, or consist of, specific neural vehicles that are uniquely identifiable and, at least in principle, distinguishable from other engrams and non-engram structures. On the hypothesis, the elicited behavioral responses can be considered indicators of the information encoded by a targeted neural ensemble. Loss-of-function and gain-of-function studies thus provide preliminary evidence for characterizing engrams as *discrete* information-bearing structures. The studies link the selective manipulation of a neural ensemble to a specific behavioral response, one purportedly not elicited by the manipulation of another ensemble. In principle, the independent manipulation of distinct engram ensembles—each carrying informational content acquired on some prior occasion—is expected to produce specific and clearly distinguishable behavioral responses.

Yet, given the nature of these studies, caution is warranted. There are two immediate concerns. The first pertains to the attribution of contentful states to rodents performing associative conditioning tasks.[23] Even a brief look at the literature reveals a controversy about whether associations, of the kinds formed in such tasks, are content-bearing. Contra much traditional wisdom, some recent accounts have indeed provided a positive answer, characterizing associative conditioning as propositional and inferentially rich (e.g., De Houwer 2009; Gallistel & King 2010). Nevertheless, attributing representational content may require more complex states and flexible use of information of a kind engram theorists are only beginning to explore (Cao 2022). The second concern pertains to the relevance of these studies for episodic memory specifically. Engram theorists have often glossed over this issue, cutting across seemingly well-established distinctions between memory kinds.[24] It nevertheless remains open whether there are different kinds of engrams and how exactly conditioning studies bear on EMF. Conservatively, we should treat them as providing one of several *converging* evidential threads, to be assessed with the other evidence presented

---

[22] The plausible dissimilarity between optogenetic reactivation and natural retrieval has been at the heart of the controversy over so-called "silent engrams", retrievable only artificially (Ryan et al. 2015).

[23] I return to the general problem of content in section 4.

[24] There is a reason for this. As Najenson (2023) emphasizes, many engram theorists have pursued a *general* theory of information storage, which has the potential to unify memory research in the neurosciences.

here. Together, the hope is, we get a relatively compelling picture of engrams as information-bearing causally relevant neural structures.

This picture, nevertheless, does not entail that the content of an engram remains invariant over time.[25] Indeed, some of the studies examined above arguably involve content change—e.g., the formation of a novel association between a fear-inducing shock and a previously neutral context. (See Robins 2016a for discussion.) The point is brought home by De Sousa et al. (2019), who showed that high-frequency stimulation of engram ensembles in the retrosplenial cortex in mice, during sleep or light anesthesia, leads to the generalization of a fear response—typically associated with processes of systems consolidation. The result dovetails nicely with the steadily accumulating evidence for content transformation during systems consolidation and reconsolidation (Winocur & Moscovitch 2011; Dudai et al. 2015). Such transformation may involve loss of content (e.g., via semanticization or schematization) as well as, arguably, incorporation of new content (e.g., from testimony or new perceptual experience).

Perhaps more counterintuitively, the local hypothesis also does not entail that the structural basis or location of an individual engram must remain invariant over time (Robins 2020). That engrams are moving targets, undergoing a variety of changes from encoding to retrieval, is the received view in the psychology of memory, one seemingly shared by engram theorists (see, e.g., Josselyn et al. 2015; Tonegawa et al. 2018). Intriguingly, a recent optogenetic study by Refaeli et al. (2023) illustrates that the neural circuits supporting a contextual fear memory in mice do reorganize over time, providing a preliminary glimpse of the moving target. The study shows that neural ensembles in cortical structures, such as the anterior cingulate cortex, become increasingly important for remote memory retrieval. *However*, a group of ensemble neurons in the dorsal hippocampus (dCA1), active during encoding, remain critical for retrieval at any time. This result, which anticipates the argument of the next section, draws attention to two significant points. First, while the structural basis of an engram need not be invariant over time, this does not entail that *all* of its structural components may undergo changes. As the study by Refaeli et al. (2023) illustrates, an invariant "core" may remain critical for memory retrieval and referential stability. The proposal developed in the next section will indeed appeal to such a core. Second, stability of reference may turn out to require some kind of *functional* stability of engram-files, compatible with significant changes in neural structure over time. What such functional stability involves and how it can be maintained, however, remain difficult open questions (cf. Robins 2020).

---

[25] More generally, transmissionism (the view that content is stored from encoding to retrieval) does not entail preservationism (the view that a memory may not include "new" content). See Michaelian & Robins (2018).

3.5. Structural Complexity and Direct Access

Having argued that the local hypothesis does not entail either content invariance, I now turn to the issue of referential stability in the face of content change. I focus on the structural complexity of engrams, hypothesized to make such stability possible. I thus examine the hippocampal indexing theory, illustrating the way in which it may point to the file-like complexity of engrams. On the emerging picture, engrams may indeed be accessible directly, a hypothesis receiving some support from evidence that event boundaries constitute access points for episodic memory retrieval.

Mental files are hypothesized to guide reference to individual entities and store information about them. On the standard portrayal, this functional role is afforded by the files' structural complexity. Theorists like Recanati (2012) thus routinely talk about *opening* particular files, storing information *in* them, and reactivating them on a later occasion to read off or potentially alter the stored information. Crucially, files retain their identity even when the informational contents they contain undergo change. The structural complexity thus affords stability of reference, making room for misattribution of properties, and even attribution of contradictory properties, to a referent. The importance of these latter phenomena to episodic memory is obvious. Indeed, much of the work in the constructivist tradition has simply taken for granted that referential stability is maintained as the informational content of a memory changes.[26] Moreover, there appears to be a class of memory errors—errors of *misremembering*—characterized by the misattribution of properties to an event successfully referred to (Robins 2016b). Finally, a memory for an event can clash with knowledge about it acquired in a different way; e.g., via testimony (Mazzoni et al. 2010).

Yet, it remains unclear whether and how engrams can be structurally complex in the relevant way, affording *direct* access and thus referential stability during content change. Indeed, skepticism about the idea has been expressed by a number of prominent theorists. The source of such skepticism is relatively straightforward and touches again upon the issue of discreteness. Ex hypothesi, engram formation involves the modification of synaptic strengths or connectivity in neural ensembles, a view often characterized as reflecting a consensus in the neuroscience of memory (see, e.g., Poo et al. 2016).[27] Yet, if this hypothesis

---

[26] In a highly characteristic example, Rubin et al. (2004) investigated the stability of subjects' ratings of various aspects of their autobiographical memories. The authors instruct the authors to recall "the *same* 20 autobiographical memories at two sessions separated by two weeks" (2004, p. 715), assuming throughout that sameness of reference is maintained despite (small) changes in content. Examples of this kind abound in the experimental literature and should give some pause to methodological naturalists keen on doing away with the notion of (singular) reference.

[27] It is worth noting, however, that *molecular* models of memory, have recently gained in popularity (Langille & Gallistel 2020; Gold & Glanzman 2021). At least in principle, these models can accommodate discrete engrams more easily. Future work will reveal how promising molecular models are and how they relate to the dominant synaptic models (Colaço & Najenson 2023).

is true—the skeptical argument goes—then engrams cannot be genuinely discrete. As an organism undergoes multiple experiences, these will affect the strength of connections—or alternatively: connectivity— in relevant ensembles. While these changes may be lasting, the effect of any *particular* experience will wash out over time (Sutton 1998; Robins 2016c). Engrams in such ensembles—if they do exist[28]—will not be uniquely identifiable structures with distinct causal histories and will thus be incapable of sustaining reference-securing chains. Rather, they will be *superposed* in the same sets of synaptic connections. Superpositional models, to borrow Sutton's (1998, p. 8) apt phrase, have "problems with *sameness*"; so, they cannot maintain sameness of reference as content changes. However referential stability in memory is secured, it is not via file-like neural structures.

This conclusion is premature. To forestall its hasty acceptance, it's worth having a brief second look at the notion of superposition. Unlike discreteness, which is a feature of representational schemes, superposition concerns the overlap in physical resources used by different representations. Importantly, superposition comes in degrees. Two event representations are *fully* superposed if the resources used to represent event 1 are coextensive with those used to represent event 2 (Van Gelder 1991; Clark 1993). When two representations are fully superposed, indeed, they are effectively merged into a single representation. In most real networks, however, representations of distinct events will be only *partially* superposed, with particular features represented in overlapping regions (e.g., by the same sets of synaptic connections). Now, there is an obvious sense in which partially superposed representations could be characterized as non-discrete, constituted as they are by overlapping neural resources. Crucially, however, such representations can still be discrete in *our*, more minimal, sense. They will be so iff they function as distinct states, uniquely identifiable by the representing system, which uses them to discriminate its relevant inputs. In such a scheme, there will be "gaps" between representations: two adjacent representations (according to some relevant ordering) will be partially superposed, yet there will be no representation between them. Engrams can thus share representational resources while remaining functionally distinct.

With this in mind, we can turn to the hippocampal indexing theory, which can be seen as providing a picture of engrams as discrete, yet partially superposed, representations. According to the theory, the hippocampus contributes to episodic memory formation by storing an *index* of the spatiotemporal pattern of (neo)cortical activity generated by a particular episode (Teyler & DiScenna 1985, 1986; Teyler & Rudy 2007). As an organism undergoes an experience, a variety of cortical areas are engaged; visual, auditory, tactile etc. Crucially, a population of active hippocampal neurons "binds" the patterns of neural activity

---

[28] Indeed, theorists characterizing engrams as *necessarily* discrete will have a hard time treating these models as anything but eliminativist about engrams. See Sutton & O'Brien (2023) for one attempt at avoiding this conclusion.

across these distributed areas—generating a kind of index for the particular experience. The subsequent reactivation of the index enables reinstatement of the cortical activity linked at encoding. Such reinstatement (e.g., during sleep) promotes systems consolidation, strengthening hippocampal-neocortical connections. More to the point, it subserves recall of *particular* experiences from the organism's past—with the hippocampal index securing the uniqueness of a memory. There is now a lot of experimental evidence for hippocampal indexing (for a review, see Goode et al. 2020). It includes the finding that optogenetic activation of a very small subset (e.g., 2-3%) of experience-tagged cells, in the dentate gyrus and CA1 region, triggers the behavioral expression of a memory (Liu et al. 2012; Ryan et al. 2015). More directly, the inhibition of hippocampal activity suppresses the reinstatement of cortical neurons active during encoding, thereby suppressing retrieval (Kitamura et al. 2017; Lee et al. 2023).

The indexing theory is, *in principle,* agnostic about the representational profiles of hippocampal neurons (see Goode et al. 2020). On the received view, however, hippocampal indices carry information about neocortical activity generated by particular events, yet *not* about the events themselves. As Teyler & Rudy (2007) put it in their influential treatment, "the hippocampus itself does not contain the content of an experience but it does provide an index that allows the content to be retrieved" (p. 1158). Event information is neocortically processed and stored (Teyler & DiScenna 1985, 1986; Teyler & Rudy 2007). We can thus think of indices as sets of *pointers* to cortically stored representations of event features. Along these lines, Fayyaz et al. (2022) propose a computational model of episodic retrieval in which hippocampal neurons do not represent properties of the input but rather function as pointers directed to perceptual and semantic elements of a representation stored in cortical areas. Reactivation of the indices triggers the reinstatement of this representation—and a complementary process of semantic completion—and is thus necessary for successful retrieval. Teyler & Rudy (2007), and a number of their theoretical successors, identified the engram *with* the hippocampal index. Yet, this conception is neither necessitated by the details of the models nor properly aligned with the functional profile of engrams as characterized in 3.1. As we saw there, an engram is a lasting neural change resulting from a specific learning experience and hypothesized to carry content about it. The formation of a hippocampal index, while crucial for maintaining engram discreteness, does not exhaust the changes due to a learning experience. It is rather only a component in a wider hippocampal-cortical network whose disposition to reinstate a relevant pattern of activity is altered by the learning experience (cf. De Brigard 2024a). It is this network that houses content-carrying (components of) event representations, capable of existing in a dormant state.[29]

We are thus better off characterized the engram as structurally *complex,* consisting of a unique hippocampal index *and* cortically stored representational components. The index secures an engram's

---

[29] I return to this issue in the following section.

distinctness and identity over time, despite partial superposition, with different engrams sharing cortical representational resources. As long as it remains critical for recall—as the evidence and the models above suggest—its tokening can secure stability of reference even as the informational content of a memory changes. A structurally complex engram of this kind will be discrete in the relevant sense and will have a unique causal history, formed upon a particular experience and causally operative in recalling it. If the model is on the right track, hippocampal indices may function as atomic pointer-like components, along the lines of some proposals by files theorists (Recanati 2012; Green & Quilty-Dunn 2021). On this view, the index-triggered reactivation of a distributed hippocampal-cortical representation functionally corresponds to "pulling the file from storage" to read off its contents.

In this context, the recent evidence that information about event boundaries may play a role in indexing is particularly intriguing. Single neurons in the human hippocampus, predictive of memory performance, have been shown to be responsive to event boundaries (Zheng et al. 2022). This finding is paralleled by the discovery of hippocampal CA1 neurons in rodents, coding for discrete events in experience (e.g., Sun et al. 2020). Relatedly, a recent study has demonstrated a dissociation of spatial coding and contextual indexing, with specific ensembles of CA1 neurons, characterized by the expression of the immediate early gene c-FOS, capable of reinstating context-appropriate behavior but showing low spatial specificity (Tanaka et al. 2018). In a working model of episodic recollection, Ross & Easton (2022) appeal to such evidence, hypothesizing that subsets of CA1 and c-FOS cells function as hippocampal indices, preferentially responsive to event boundaries and coordinating reinstatement of distributed cortical activity patterns.

On the emerging picture, we can characterize a form of direct engram access, which resembles the one envisioned by mental files theorists. *Index-addressable* engrams can be accessed by activation of their hippocampal index and not via their cortically-stored informational entries (Tanaka et al. 2018; Goode et al. 2020). If hippocampal indexing mechanisms do indeed process information about event boundaries, then evidence that these constitute access points for episodic memory retrieval would be of particular importance. A recent study by Michelmann et al. (2023) points in this direction. After having subjects watch movies in a naturalistic setting, the experimenters asked them questions that required orientation first to a specific moment A and then to a later moment B in the movie. This allowed Michelmann et al. to measure the time it took subjects to "get" from A to B in memory. They found that the number of distinct events in the relevant segment and the distance from B to the previous event boundary were most significant in explaining the subjects' scanning time (more so than segment duration). This illustrates that memory search—at least in a continuous experience—is not primarily based on semantic relatedness but rather on underlying event structure, with event boundaries constituting "natural" access points for retrieval. The finding fits nicely with the idea that event segmentation affects the organization of information in long-term memory as well as with recent neural network models highlighting the computational advantages of selectively encoding episodic memories at the end of events (Lu et al. 2022).

Results of this kind provide preliminary evidence for the index-addressability of engrams. Yet, here again, we should proceed with care. Given the lack of agreement about what (if anything) indexing neurons represent, it would be premature to conclude that, even in the highlighted cases, engrams' informational contents do not play active roles in memory search and access. It is, moreover, an open question how the results fit in the larger picture. While theorists have long hypothesized that episodic memories are *both* index- and content- addressable, combining the different forms of access in a comprehensive, and empirically respectable, account remains a major challenge (Anderson et al. 1998; Frankland et al. 2019; Kahana 2020). The presented evidence thus provides tentative support for the direct access hypothesis.

I have characterized the deployment of an engram in recall as involving cortical reinstatement. It is tempting to see such process as involving simple reactivation of cortical neural populations linked at encoding. Such a view, however, would be incompatible with the wealth of evidence pointing to the systematic transformation of neural activity from perception to memory retrieval (Favila et al. 2020). The indexing theory provides a principled reason why—even putting aside considerations about function—such changes are likely to occur. As only a limited information about an event is initially encoded, and such information is often unavailable at retrieval (for a number of obvious reasons), index-driven cortical reinstatement is essentially reconstructive. It relies on processes of semantic completion and regularly recruits neural populations not involved during encoding (Rolls et al. 2013; De Brigard 2024a). Pulling an engram-file from storage to read off its contents is thus often accompanied by writing something new in it, a design feature which may indeed facilitate the maintenance of a memory's relevance (Lee 2009) or maximize the rememberer's predictive prowess (Schacter et al. 2012).[30] Before moving on, we should highlight an important issue concerning these models. If hippocampal neurons index neocortical activity *and* such activity varies across retrievals, then the index itself cannot remain structurally invariant over time. In such circumstances, the functional identity of engrams—and, consequently, referential stability—will likely be secured by the causal-informational links connecting the different structural bases.[31] This

---

[30] This admittedly counterintuitive model portrays engrams as complex structures with informational entries that are regularly altered, often in unpredictable ways. Andonovski (2021b) illustrates the point, using Plato's aviary metaphor: "[A] lot of interesting stuff seems to happen to representations "in memory": birds flock and shift allegiances, they alter their strength and appearance, morph with other birds, and sometimes...they simply disappear" (p. 237).

[31] This proposal aligns well with the *multiple trace theory,* which posits that repeated retrievals lead to the formation of multiple structurally distinct states (individuated as distinct traces), causally active in subsequent acts of recall (Nadel & Moscovitch 1997). Importantly, the theory does not deny the existence of uninterrupted causal-informational chains between memories and past experiences, or the necessity of engrams for episodic recall, so it is in principle compatible with EMF. I am thankful to a reviewer for highlighting this important issue.

proposal is compatible with the time-dependent reorganization of engram access (Teyler & Rudy 2007; Goode et al. 2020) as well as with the transformation of some engrams into so-called "encyclopedia entries", not individuated in terms of specific ER relations to referent events (Recanati 2012).

4. Engrams and Causal-Representationalist Theories

4.1. Engram-files: Between Replay and Props

The last few sections compose an evidential mosaic, with pieces of suggestive, but preliminary, evidence from a variety of disciplines, theoretical frameworks, and experimental paradigms. The picture that emerges portrays engrams as information-bearing physical structures, formed upon individual experiences of events and causally involved in their subsequent recall. Plausibly discrete, they exist in a dormant state and are behaviorally expressed through interaction with retrieval cues. Intriguingly, engrams appear to exhibit a file-like structural complexity. A complex engram, consisting of a hippocampal index and a distributed cortical representation of an event's properties, may afford direct access and, hence, referential stability during content change. I have argued that the deployment of engrams is constitutive of the capacity for singular reference in episodically remembering particular events. Engrams sustain reference-securing epistemically rewarding relations to events in the organism's past, permitting information flow from past to present.

This characterization aims to combine the virtues of two competing images of engrams, recently brought out by Langland-Hassan (2022). The first portrays engrams as direct recordings of past experiences, whose reactivation—i.e., *replay*—drives episodic recall. Replay-anchoring engrams are "monogamous", each enabling the remembering of, and preserving information about, just one event. The second, in contrast, portrays engrams as *prop-like*: reusable bits of information, which may have been acquired on individual occasions but which can be flexibly recombined and deployed in remembering many distinct events. Prop-like engrams are thus "promiscuous". As discrete informational structures, engram-files are monogamous, enabling the remembering of only those events that were causally involved in their formation (even if some of these events are temporally protracted and quite complex.) Remembering does involve reactivation of cortical, and potentially subcortical, areas involved in the original experience. Yet, as we saw in the last section, index-driven cortical reinstatement does not resemble simple replay and should rather be seen as reconstructive in nature, resulting in occasional changes in memory contents. On the emerging picture, indeed, information acquired on an individual occasion, and linked to a particular hippocampal index, can nevertheless be deployed in remembering other events—kind of like a prop. Engram-files may be monogamous but their informational entries have an exciting amorous life.

The EMF view resembles two popular accounts of engrams, developed in the recent literature. In the remainder of this section, I'll highlight the important similarities and differences. The first, *minimalist,* account can be traced to the original indexing theory. Developing it, Teyler & DiScenna (1986) proposed a "division of labor" between neurocognitive systems. The hippocampus stores—and *retains*—indices of spatiotemporal activity in the neocortex, which in turn represent features of external events. A key assumption was that the neocortical patterns of activity were transient and not accompanied by event-specific modification of neocortical connectivity—at least in the short run. Since, *ex hypothesi,* engrams are event-specific, Teyler and colleagues proposed to identify the engram with the hippocampal index. The result was a surprising commitment to a view of engrams as discrete but contentless neural structures (see, e.g., Teyler & Rudy 2007, p. 1163). Along similar lines, Werning (2020) has more recently argued that a "minimal trace"—discrete and causally operative in recall—carries sparse information (e.g., about neural firing sequences) but not representational content.

The minimalist account is anchored on a perceived tension between experience-dependence and contentfulness: if engrams exhibit the first property, they cannot exhibit the second. The "contentless engram" solution offers an intriguing conceptual revision, pushing engram theory strongly in a constructivist direction. Future work will determine whether such revision is genuinely warranted. Here, I want to heighten the contrast with EMF by highlighting two important points—one conceptual, one empirical. First, the endorsement of an indexing theory does not compel the identification of the engram with the hippocampal index. As I have argued, indexing can effectively "discretize" widely distributed and superposed representations. Structurally complex engrams may be both experience-dependent (i.e., event specific) *and* contentful, rendering the tension between the two properties only apparent. Second, recent evidence of rapid cortical learning and plasticity has the potential to undermine the key minimalist assumption. In a representative study, Kitamura et al. (2017) found prefrontal "engram cells", critical for contextual fear memory, to be rapidly generated during initial learning, through inputs from both the hippocampus and the amygdala. A number of studies have similarly revealed rapid event-specific cortical modifications (Cowansage et al. 2014; Refaeli et al. 2023; for discussion, see Hebscher et al. 2019). While this work is in its nascent stages, it does point to widespread and seemingly endurant neural changes due to singular learning experiences. The wager of the EMF view is that a maximalist account of engrams—i.e., one which characterizes them as both experience-dependent *and* contentful—will be in a better position to accommodate such evidence than a minimalist one.

The second popular account, developed by De Brigard (2014, 2024a), characterizes engrams in *dispositionalist* terms. De Brigard endorses a revised version of the indexing theory, describing the hippocampal index as encoding a set of instructions for the reactivation of a cortical pattern of activity.

Unlike the minimalist, however, he does not identify the engram with the index. Rather, he thinks of it as a dispositional property of the neural network:

> [W]hen one experiences a certain event during encoding, the experienced content is instantiated in a particular representational vehicle, in the form of a hippocampal-neocortical network... When [a right] cue is presented in the retrieval context, the coactivation among units of the network starts to propagate toward the hippocampal index, which does not contain explicit contents but rather the conditional instructions to reactivate the rest of the pattern of activity. This, I suggest, is the right way to understand what [an engram] is: the dispositional property of a neural network to reinstate the state it was in, during encoding, at the time of retrieval (2024, p. 16).

Such a network may be rapidly modified by singular learning experiences, yet event-specific modifications are unlikely to endure, as the network gets constantly redeployed in a variety of tasks—and thus further transformed (De Brigard 2014, p. 411).

It is not clear what to make of this dispositionalist account. A key question is whether the hypothesized dispositions have categorical causal bases that exhibit the typical functional properties of engrams.[32] If the answer is positive, of course, the account is compatible with EMF. Charity dictates, however, that De Brigard favors a negative answer. This renders the dispositionalist account highly revisionist. Engrams are traditionally invoked to causally *explain* the ability of organisms to reliably represent previously experienced events. On the indexing theory, they do so by playing a critical causal role in the reinstatement of cortical patterns of activity (i.e., the vehicles of "retrieved" memory representations). By identifying engrams *with* dispositional properties of networks to reinstate specific activation patterns—properties thus individuated in terms of their manifestations—De Brigard moves decisively away from a causal-explanatory account. On the resultant picture, engrams do not causally explain acts of remembering; rather, they are constitutively linked to them. Even theorists tolerant of extensive concept revision would be hard-pressed to characterize the dispositionalist account as anything but eliminativist about engrams. Whether such eliminativism is warranted is, of course, an open empirical question. Here, I have tried to provide some reasons to resist it. On the EMF view, an engram owes its particular structure and function to a unique learning experience undergone by the organism. Engrams may exist in dormant states, yet they are—at least in principle—uniquely identifiable and independently manipulable. Such targeted manipulations, as we have seen, appear content-sensitive as they produce the behavioral effects expected on the assumption that engrams carry specific representational contents. Hence, while a variety of causes

---

[32] There is considerable controversy in the literature concerning the distinction between dispositional and categorical properties. The key intuition is that, unlike categorical properties, dispositional properties are *essentially* tendencies to produce certain effects. Despite this, philosophers disagree about whether "dispositional" and "categorical" are contraries as well as, among other things, about whether there are properties with dispositional essences. For discussions, see Bird (2007) and Choi (2011). Thanks to a reviewer for encouraging me to include this clarification.

may play important roles in supporting the dispositional property of a hippocampal-cortical network to reinstate a previous state, there is no *principled* reason to think that the disposition does not have an engram-like categorical basis.

Before moving on, we need to address an important point. The characterization of engrams as mental files may strike some theorists as a kind of category mistake. For Recanati (2012), as for many others, mental files are conceptual *constituents* of thought; they are “the mental counterparts of singular terms” (p. 35). Engrams, in contrast, are typically seen as *sub*-personal structures manipulated by memory systems. This is a thorny issue, even for those who fully take to heart that local hypotheses are unlikely to provide smooth reductions. Conservatively, we can treat engrams as *proto*-files (Recanati 2012), non-conceptual structures that make possible—perhaps by providing basic “bundling principles” for the integration of information (Campbell 2002)—the subsequent deployment of singular conceptual representations in memory. Yet, there are reasons to be more optimistic. As we have seen, engrams exhibit some of the hallmark features of conceptual representations, with indices potentially guiding access to stored information about referent events. Deployed in recall, they may thus function as singular term constituents in the “language” of episodic thought (Quilty-Dunn et al. 2023; Mahr & Schacter 2023). Nevertheless, this picture will likely require giving up the introspective transparency of reference in memory. While the investigation of this issue is at its infancy, the matching of behavioral and neural reinstatement in experimental contexts suggests that *some* form of transparency may be preserved, despite the absence of direct introspective access to engram vehicles. Transparency may indeed not require such access (Murez 2023).

In the final section, I build on the EMF view to examine the prospects for the development of a causal-representational theory of memory.

4.2. Toward a Causal-Representationalist Theory

Martin & Deutscher's (1966) causal-representationalist theory posited engrams from the proverbial armchair, characterizing the notion of engram as an essential component of commonsense conceptions of memory. The theory was rightly criticized for "dictating" to memory science what to discover (Zemach 1983) and subsequently, indeed, for being incompatible with it (Michaelian 2016). In this paper, I have taken the other—*naturalist*—route, treating my proposal as a working empirical hypothesis about the role engrams play in sustaining reference-securing relations to past events and thereby anchoring episodic singular thought. For such hypothesis to be at the basis of a causal-representational (C-R) theory, further work is required—conceptual and empirical. It includes specification of the nature and objectives of such a theory as well as integration of evidence about engrams into a more comprehensive picture of memory processes and phenomena. While the bulk of this work has to be done on another occasion, I use the last section to offer a provisional sketch. I focus on the nature of the theory, outlining its major features, open questions, and challenges.

A naturalist philosophical theory aims to develop a general, high-level picture of a target phenomenon by engaging with productive research programs in the sciences. Such a theory will have a characteristically broad scope, integrating insights from a variety of sources and bringing them into contact with issues of philosophical interest. As Michaelian (2016, p. 3) programmatically points out, a naturalist theory of memory "aims 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." I have attempted to show that the vision of memory, underlying a variety of new developments in the study of engrams and memory processes more generally, is a familiar one. It sees memory as an essentially *retentive* capacity, bringing organisms into epistemic contact with the past—albeit one mediated by reconstructive and transformational processes. The contact is made possible by the formation, maintenance, and retrieval of discrete, file-like informational structures.

Yet, the focus of the theory will be specifically on *episodic* memory, a capacity thought to be underlaid by a dedicated, functionally dissociable, memory system. As O'Sullivan & Ryan (2024) forcefully argue, theorizing about engrams has to be accompanied by a characterization of the computational problems engram-mongering systems, and their constituent processes, aim to solve. A proper understanding of mnemic retention and reference thus requires knowledge of the normal functioning of the episodic system. At a minimum, this would involve understanding of the procedures that underlie the deployment of engrams in episodic recall, characterization of the ways in which information retention and change are balanced as well as specification of the multiple uses to which such information is seemingly put to (e.g., in imagination or counterfactual thought). A system-centric C-R theory will, accordingly, examine episodic memory's specialized processes and states and the principles that guide its interactions with other (memory) systems in the production of introspectively and behaviorally accessible phenomena. It will *not* entertain extra-theoretical criteria of "mnemicity" or "episodicity" or specify necessary conditions token states must satisfy to *be* episodic memories (Michaelian & Sutton 2017). As the episodic system is imperfectly reliable, some prototypical memory states will not have the "appropriate" etiological profile and will thus misrepresent their targets.[33] At least in principle, these will not provide counterexamples to the theory (cf. Khalidi 2023).

---

[33] The target of a memory representation is whatever it is *supposed to* represent on the particular occasion (Cummins 1996). A memory representation can *mis*represent, in which case its target is distinct from its content/referent. Thanks to an anonymous reviewer for raising this issue.

I have characterized engrams as playing a dual functional role: securing reference to, and carrying information about, previously experienced events. Thus characterized, engrams function as "first-level explainers" (O'Sullivan & Ryan 2024), directly accounting for features of behaviorally manifested memory phenomena. Indeed, the emerging picture is strongly consilient with Tulving's (1983) classic theory of episodic memory as a system with a proprietary store and retrieval procedures. An engram-file enables reliable recollection of an event, providing a kind of "cognitive blueprint" specifying the conditions under which such recollection will occur (Tulving & Watkins 1975). Sustaining a causal-information relation to an event, it affords direct mnemic reference to it. Such directness may be closely linked to a memory's mode of presentation. As we saw, episodic memories have been thought to have a distinctive mode of presentation, with remembered events presented in thought *as* previously experienced—a mode hypothesized to correspond, in humans at least, to an autonoetic consciousness (Tulving 1985; Sant'Anna et al. 2023). Despite these similarities, the episodic memory system is likely to be more "inferentially capable" than Tulving originally suspected.

I have appealed to the theory of mental files to highlight the tight link between the causal and informational aspect of a memory's dependence on prior mental states. Such dependence, as causalists have insisted, has to be undergirded by a causal chain linking perceptual experience and encoding to subsequent recall. As I have tried to illustrate, this chain is unlikely to be brutely causal and, at a minimum, involves a variety of plausibly inferential, information-mongering processes. Relatedly, much of the evidence for the causal relevance of engrams comes from the variable behavioral expression of a memory due to the manipulation of an ensemble hypothesized to carry information about a positively or negatively valenced experience. Once we accept the causal model for memory, to reiterate Martin & Deutscher's (1966, p. 189) claim, we must also accept the existence of an information-bearing engram. It is worth noting in this context that the adoption of this causalist perspective casts little light on the truly difficult problem of memory retrieval. Even if every experience an organism undergoes results in a unique engram, the system still has to locate the right engram, in the appropriate epistemic context. How this is accomplished remains a core question for future theoretical and empirical research (Frankland et al. 2019; Andonovski 2021b).

How much change in the contents of a memory a properly functioning episodic system can "tolerate"—particularly in light of the posited informational dependence of recall on encoding—also remains an open question. This question has been at the forefront of the debate between causalists and simulationists in the philosophy of memory, with the former posing various limitations on the nature and extent of allowable change (Bernecker 2008; Michaelian 2016). While the naturalist C-R theorist will have some sympathy for these proposals, they would be hesitant to adopt them from the armchair. If the local hypothesis is on the right track, then there must be *some* limits to content change, even if these vary with the context and the nature of the task. Yet, the hypothesis has to be brought into contact with prominent considerations about the importance of transformational and reconstructive processes—e.g., for the maintenance of a memory's relevance (Lee 2009). A deeper understanding of systems' functioning, its proprietary processes and uses of information, will be necessary to shed light on these issues.

A more prominent question concerns the status of engrams as *representations,* hypothesized to carry content about remembered events or experiences. This is a big topic, but it's worth outlining the approach that best reflects the methodological stance highlighted above. The naturalist C-R theorist sees the problem of content not as requiring a philosophical theory of intentionality—to be constructed independently of first-order theorizing—but rather as an invitation to examine the ways in which candidate neural states function as representations in the course of a system's natural functioning. Ramsey (2023) expresses this idea with some flair, wagering that a proper appreciation of it would lead to an eventual dissolution of the "content question":

> Instead of treating a theory of content as the avenue for understanding how brain states could function as representations, we should instead regard a theory of representation function as the avenue for making sense of content… Once we remind ourselves that representations are a *functional* sort of thing… we can begin to see that these purportedly problematic features of intentionality are actually somewhat mundane features of many functional kinds (p. 9).

In this spirit, philosophers have become increasingly engaged with the nitty-gritty details behind content attributions in the mind sciences. In a notable contribution, Cao (2022) argues that the neuroscientific practice of attributing content to patterns of neural activity is justified by their re-identifiability and causal relevance for the production of pertinent behavioral effects—as manifested by their effective manipulation (see also Gładziejewski & Miłkowski 2017; Piccinini 2022). In the past sections, I provided examples of such "content-sensitive" manipulations, aiming to integrate them in a more comprehensive picture of engram-mongering memory systems. While there are residual concerns about the details (see 3.4.), these do not constitute principled obstacles to a representationalist approach. Future work, theoretical and experimental, will ascertain whether, and under which conditions, memory theorists are entitled to appeal to contentful engrams.

I end with a few words about cognitive significance. On Tulving's (1985) influential model, the utilization of engram information in memory corresponds to the way the event is "given" in thought—an episodic mode of presentation. In humans at least, this is reflected in the autonoetically-flavored recollective experience and the disposition of subjects to characterize remembered events as ones they have previously experienced, typically paired with a tendency to assert epistemic authority with respect to them. Four decades later, the extensive research on the phenomenological and behavioral signatures of episodic recollection remains poorly integrated with what we know about engrams and event cognition. As engram technology and our knowledge of memory processes mature, we should expect bolder attempts at integration. These may illuminate both the cognitive role of engrams and the ways in which they contribute to the phenomenology of episodic recollection.

5. Conclusion

I have used the theory of mental files to introduce and preliminarily defend an empirical hypothesis about the nature and function of engrams. Bringing together evidence from different research programs in the memory sciences, I characterized engrams as discrete informational structures, formed upon prototypically perceptual experiences of events. Plausibly exhibiting a file-like structural complexity, engrams afford transmission of content from such experiences to subsequent states of recall and are causally involved in their production. The evidence, I argued, should make us reasonably confident about the capability of engrams to anchor episodic singular reference and, more generally, about the prospects of naturalist causal-representational theories.

This approach is likely to strike many as misguided or overly optimistic. To mental files theorists, the examination of the nitty-gritty details of memory processing may seem unnecessary or insufficiently respectful of the autonomy of philosophical theorizing. To engram theorists, the focus on reference may appear like an introduction of an external thematic framework, not properly aligned with neuroscientific research. We should be sensitive to such criticisms but remain confident about the value of integrative proposals. Philosophical theorizing about reference, while perhaps autonomous in some sense, cannot be divorced from the study of the neurocognitive mechanisms that allow organisms to relate to the past in an epistemically rewarding way. And, in the long run, psychological and neuroscientific evidence about such mechanisms will have to be brought into contact with issues about reference that have traditionally piqued the interest of philosophers. While it may be too early to attempt this, we should be optimistic; if not about the proposed integration, then at least about the likelihood that we will learn something from our failure.

Acknowledgements

I thank the editors of the topical collection "Reference and Remembering", and particularly James Openshaw, for the invitation, support, and patience. Thanks also to my CPM colleagues, to Jonathan Najenson, Sarah Robins, Michael Murez, François Recanati, Markus Werning, two anonymous reviewers, and audiences at "Reference in Remembering" and "Memory and Mental Files" at CPM, "Generative Episodic Memory" at Ruhr-Universität Bochum, and the 3rd Joint Meeting of SPP and ESPP in Milan. This research was funded by the European Union's Horizon research and innovation program under the Marie Sklodowska-Curie grant agreement 101062754.

Bibliography:

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

Aly, M. and Turk-Browne, N. B. (2016). Attention promotes episodic encoding by stabilizing hippocampal representations. *Proceedings of the National Academy of Sciences*, 113(4):E420{E429

Andermane, N., Joensen, B. H., & Horner, A. J. (2021). Forgetting across a hierarchy of episodic representations. *Current Opinion in Neurobiology*, 67, 50-57.

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language*, 38(4), 341-380.

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

Andonovski, N. (2021b). Memory as triage: facing up to the hard question of memory. *Review of Philosophy and Psychology*, 12(2), 227-256.

Andonovski, N. (2023). Autonoesis and the Galilean science of memory: Explanation, idealization, and the role of crucial data. *European Journal for Philosophy of Science*, 13(3), 42.

Andonovski, N., & Michaelian, K. (Forthcoming). Naturalism and simulationism in the philosophy of memory. In Hossein Khani, A., Kemp, G., Rezaei, H. S., & Amiriara, H. (Eds.), *Naturalism and its challenges*. Routledge.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (Vol. 2, pp. 89-195). Academic press.

Bach, K. (1997). *Thought and reference.* Oxford: Clarendon Press

Bailey, H. R., Kurby, C. A., Sargent, J. Q., & Zacks, J. M. (2017). Attentional focus affects how events are segmented and updated in narrative reading. *Memory & Cognition*, 45(6), 940–955.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron,* 95(3), 709–721.e5.

Barry, D. N., & Maguire, E. A. (2019). Remote memory and the hippocampus: A constructive critique. *Trends in cognitive sciences,* 23(2), 128-142.

Bernecker, S. (2008). *The metaphysics of memory.* Springer Science & Business Media

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

Bird, A. (2007). *Nature's metaphysics: Laws and properties*. OUP Oxford.

Boltz, M. (1992). Temporal accent structure and the remembering of filmed narratives. *Journal of Experimental Psychology: Human Perception and Performance,* 18(1), 90.

Brette, R. (2019). Is coding a relevant metaphor for the brain?. *Behavioral and Brain Sciences*, 42, e215.

Campbell, J. (2002). *Reference and consciousness.* Oxford University Press.

Cao, R. (2022). Putting representations to use. *Synthese,* 200(2), 151.

Choi, S. (2012). Intrinsic finks and dispositional/categorical distinction. *Noûs*, *46*(2), 289-325.

Clark, A. (1993). *Associative engines: Connectionism, concepts, and representational change*. MIT Press.

Cohn-Sheehy, B. I., Delarazan, A. I., Crivelli-Decker, J. E., Reagh, Z. M., Mundada, N. S., Yonelinas, A. P., ... & Ranganath, C. (2022). Narratives bridge the divide between distant events in episodic memory. *Memory & Cognition*, 50(3), 478-494.

Colaço, D., & Najenson, J. (2023). Where memory resides: Is there a rivalry between molecular and synaptic models of memory?. *Philosophy of Science*, 1-11.

Collin, S. H., Milivojevic, B., & Doeller, C. F. (2015). Memory hierarchies map onto the hippocampal long axis in humans. *Nature neuroscience,* 18(11), 1562-1564.

Cowansage, K.K, Shuman,T., Dillingham, B.C., Chang, A., Golshani, P., Mayford, M. (2014). Direct reactivation of a coherent neocortical memory of context. *Neuron* 84: 432–41.

Craik, F. I. (2020). Remembering: An activity of mind and brain. *Annual review of psychology*, 71, 1-24.

Crane, T. (2011). The singularity of singular thought. In *Aristotelian Society supplementary volume* (Vol. 85, No. 1, pp. 21-43). Oxford, UK: Oxford University Press.

Cummins, R. (1996). *Representations, targets, and attitudes*. MIT press.

Dafni-Merom, A., and Arzy, S. (2020). The radiation of autonoetic consciousness in cognitive neuroscience: a functional neuroanatomy perspective. *Neuropsychologia* 143:107477

De Brigard, F. (2014). The nature of memory traces. *Philosophy Compass*, *9*(6), 402-414.

De Brigard, F. (2024a). Simulationism and memory traces. In S. Aronowitz & L. Nadel (Eds.), *Memory, Space and Time*. Oxford: Oxford University Press.

De Brigard, F. (2024b). Episodic memory without autonoetic consciousness. *Philosophical Transactions B,* 379 (1913), 20230410.

De Freitas, J., Liverence, B. M., & Scholl, B. J. (2014). Attentional rhythm: A temporal analogue of object-based attention. *Journal of Experimental Psychology: General*, 143(1), 71.

De Sousa, A. F., Cowansage, K. K., Zutshi, I., Cardozo, L. M., Yoo, E. J., Leutgeb, S., & Mayford, M. (2019). Optogenetic reactivation of memory ensembles in the retrosplenial cortex induces systems consolidation. *Proceedings of the National Academy of Sciences*, 116(17), 8576-8581.

Deisseroth, K. (2010). Controlling the brain with light. *Scientific American*, 303(5), 48-55.

Denny, C. A., Kheirbek, M. A., Alba, E. L., Tanaka, K. F., Brachman, R. A., Laughman, K. B., ... & Hen, R. (2014). Hippocampal memory traces are differentially modulated by experience, time, and adult neurogenesis. *Neuron,* 83(1), 189-201.

Dietrich, E., & Markman, A. B. (2003). Discrete thoughts: Why cognition must use discrete representations. *Mind & Language,* 18(1), 95-119.

Draaisma, D. (2000). *Metaphors of memory: A history of ideas about the mind.* Cambridge University Press.

DuBrow, S., & Davachi, L. (2013). The influence of context boundaries on memory for the sequential order of events. *Journal of Experimental Psychology. General*, 142(4), 1277–1286.

DuBrow, S., & Davachi, L. (2016). Temporal binding within and across events. *Neurobiology of Learning and Memory*, 134, 107–114.

Dudai, Y., Karni, A., & Born, J. (2015). The consolidation and transformation of memory. *Neuron*, 88(1), 20-32.

Eichenbaum, H. (2016). Still searching for the engram. *Learning & behavior*, 44, 209-222.

Evans, G. (1982). *The varieties of reference.* Oxford: Clarendon Press

Ezzyat, Y., & Davachi, L. (2011). What constitutes an episode in episodic memory? P*sychological Science*, 22(2), 243–252.

Faber, M., & Gennari, S. P. (2015). In search of lost time: Reconstructing the unfolding of events from memory. *Cognition,* 143, 193–202

Favila, S. E., Lee, H., & Kuhl, B. A. (2020). Transforming the concept of memory reactivation. *Trends in neurosciences,* 43(12), 939-950

Fayyaz, Z., Altamimi, A., Zoellner, C., Klein, N., Wolf, O. T., Cheng, S., & Wiskott, L. (2022). A model of semantic completion in generative episodic memory. Neural Computation, 34(9), 1841-1870.

Frankland, P. W., Josselyn, S. A., & Köhler, S. (2019). The neurobiological foundation of memory retrieval. *Nature neuroscience*, 22(10), 1576-1585.

Gallistel, C. R., & King, A. P. (2010). *Memory and the computational brain: Why cognitive science will transform neuroscience*. John Wiley & Sons.

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, 16, 309-313

Geerligs, L., Gözükara, D., Oetringer, D., Campbell, K. L., van Gerven, M., & Güçlü, U. (2022). A partially nested cortical hierarchy of neural states underlies event segmentation in the human brain. *ELife,* 11, e77430

Gershman, S. J. (2023). The molecular memory code and synaptic plasticity: A synthesis. *Biosystems*, 104825.

Gładziejewski, P., & Miłkowski, M. (2017). Structural representations: Causally relevant and different from detectors. *Biology & Philosophy,* 32, 337-355.

Gold, A. R., & Glanzman, D. L. (2021). The central importance of nuclear mechanisms in the storage of memory. *Biochemical and Biophysical Research Communications,* 564, 103-113.

Goode, T. D., Tanaka, K. Z., Sahay, A., & McHugh, T. J. (2020). An integrated index: engrams, place cells, and hippocampal memory. *Neuron,* 107(5), 805-820.

Goodman, R., & Gray, A. (2022). Mental filing. *Noûs,* 56(1), 204-226.

Green, E. J. and Quilty-Dunn, J. (2021). What is an object file? *The British Journal for the Philosophy of Science*, 73(3):665–699

Guskjolen, A., Kenney, J. W., de la Parra, J., Yeung, B. R. A., Josselyn, S. A., & Frankland, P. W. (2018). Recovery of "lost" infant memories in mice. *Current Biology,* 28(14), 2283-2290.

Guskjolen, A., & Cembrowski, M. S. (2023). Engram neurons: Encoding, consolidation, retrieval, and forgetting of memory. *Molecular Psychiatry*, 1-13.

Halpern, D. J., Tubridy, S., Davachi, L., & Gureckis, T. M. (2023). Identifying causal subsequent memory effects. *Proceedings of the National Academy of Sciences*, 120(13), e2120288120.

Hansen, C., & Rey, G. (2016). Files and singular thoughts without objects or acquaintance: The prospects of Recanati's (and others') "Actualism". *Review of Philosophy and Psychology*, 7, 421-436.

Hard, B. M., Tversky, B., & Lang, D. S. (2006). Making sense of abstract events: Building event schemas. Memory & cognition, 34(6), 1221-1235.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory.* Psychology press.

Hebscher, M., Wing, E., Ryan, J., & Gilboa, A. (2019). Rapid cortical plasticity supports long-term memory formation. *Trends in Cognitive Sciences*, *23*(12), 989-1002.

Heinen, R., Bierbrauer, A., Wolf, O. T., & Axmacher, N. (2023). Representational formats of human memory traces. *Brain Structure and Function*, 1-17.

Heusser, A. C., Ezzyat, Y., Shiff, I., & Davachi, L. (2018). Perceptual boundaries cause mnemonic trade-offs between local boundary processing and across-trial associative binding. *Journal of Experimental Psychology. Learning, Memory, and Cognition,* 44(7), 1075–1090.

Horner, A. J., Bisby, J. A., Bush, D., Lin, W. J., & Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. *Nature communications,* 6(1), 7462.

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior,* 37(1), 1-20.

Jenkins, L. J. and Ranganath, C. (2010). Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. *Journal of Neuroscience*, 30(46):15558{15565.

Jeshion, R. (2010). Singular thought: Acquaintance, semantic instrumentalism, and cognitivism. In R. Jeshion (ed.). *New essays on singular thought* (pp.105-141). Oxford University Press

Joensen, B. H., Gaskell, M. G., & Horner, A. J. (2020). United we fall: All-or-none forgetting of complex episodic events. *Journal of Experimental Psychology: General,* 149(2), 230.

Josselyn, S. A., Köhler, S., & Frankland, P. W. (2015). Finding the engram. *Nature Reviews Neuroscience*, 16(9), 521-534.

Josselyn, S. A., & Tonegawa, S. (2020). Memory engrams: Recalling the past and imagining the future. *Science*, 367(6473), eaaw4325.

Kahana, M. J. (2020). Computational models of memory search. *Annual Review of Psychology*, 71, 107-138.

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

Kitamura, T., Ogawa, S. K., Roy, D. S., Okuyama, T., Morrissey, M. D., Smith, L. M., ... & Tonegawa, S. (2017). Engrams and circuits crucial for systems consolidation of a memory. *Science,* 356(6333), 73-78.

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in cognitive sciences,* 12(2), 72-79.

Langille, J. J., & Gallistel, C. R. (2020). Locating the engram: Should we look for plastic synapses or information-storing molecules?. *Neurobiology of learning and memory,* 169, 107164.

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

Lee, J. L. (2009). Reconsolidation: maintaining memory relevance. *Trends in neurosciences,* 32(8), 413-420.

Lee, J. H., Kim, W. B., Park, E. H., & Cho, J. H. (2023). Neocortical synaptic engrams for remote contextual memories. *Nature Neuroscience*, 26(2), 259-273.

Liu, X., Ramirez, S., Pang, P. T., Puryear, C. B., Govindarajan, A., Deisseroth, K., & Tonegawa, S. (2012). Optogenetic stimulation of a hippocampal engram activates fear memory recall. *Nature,* 484(7394), 381-385.

Lu Q., Hasson U., Norman K. A. (2022). A neural network model of when to retrieve and encode episodic memories. *eLife,* 11, Article e74445

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

Mahr, J. B., Greene, J. D., & Schacter, D. L. (2021). A long time ago in a galaxy far, far away: How temporal are episodic contents?. *Consciousness and cognition*, 96, 103224

Mahr, J. B., & Schacter, D. L. (2023). A language of episodic thought?. *Behavioral and Brain Sciences*, 46, e283.

Maley, C. J. (2011). Analog and digital, continuous and discrete. *Philosophical Studies*, *155*, 117-131.

Martin, C. B., & Deutscher, M. (1966). Remembering. *The Philosophical Review*, 75(2), 161-196.

Mazzoni, G., Scoboria, A., & Harvey, L. (2010). Nonbelieved memories. *Psychological Science*, 21(9), 1334-1340.

McKenzie, S., Frank, A. J., Kinsky, N. R., Porter, B., Rivière, P. D., & Eichenbaum, H. (2014). Hippocampal representation of related and opposing memories develop within distinct. Hierarchically Organized Neural Schemas. *Neuron*, 83(1), 202–215.

Michaelian, K. (2011). Generative memory. *Philosophical psychology*, *24*(3), 323-342.

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

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., & Sutton, J. (2017). Memory. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosoph*y. Retrieved from https://plato.stanford.edu/archives/ sum2017/entries/memory/

Michelmann, S., Hasson, U., & Norman, K. A. (2023). Evidence that event boundaries are access points for memory retrieval. *Psychological Science*, 34(3), 326-344.

Moscovitch, M. (2007). Memory: Why the engram is elusive. In H. L. Roediger III, Y. Dudai, & S. M. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 17–21). Oxford: Oxford University Press.

Murez, M., & Smortchkova, J. (2014). Singular thought: Object-files, person-files, and the sortal PERSON. *Topics in cognitive science*, 6(4), 632-646.

Murez, M., Smortchkova, J., & Strickland, B. (2020). The mental files theory of singular thought: A psychological perspective. In R. Goodman, J. Genone, & N. Kroll (Eds.), *Singular thought and mental files* (pp. 107–142). Oxford University Press

Murez, M. (2023). The transparency of mental vehicles. *Noûs.* Published online Nov 14, 2023. https://doi.org/10.1111/nous.12483

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current opinion in neurobiology,* 7(2), 217-227.

Najenson, J. (2021). What have we learned about the engram?. *Synthese*, 199(3-4), 9581-9601.

Najenson, J. (2023). LTP Revisited: Reconsidering the Explanatory Power of Synaptic Efficacy. *Review of Philosophy and Psychology*, 1-26

Nelson, M. (2002). Descriptivism defended. *Noûs*, 36(3), 408-435.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of personality and social psychology*, *28*(1), 28.

O'Sullivan, F. M., & Ryan, T. J. (2024). If engrams are the answer, what is the question? In Gräff, J. & Ramirez, S. (eds.) *Engrams: A Window into the Memory Trace* (pp. 273-302). Springer: Advances in Neurobiology 38.

Openshaw, J. (2023). (In defence of) preservationism and the previous awareness condition: What is a theory of remembering, anyway?. *Philosophical Perspectives*, *37*(1), 290-307.

Perrin, D. (2018). A case for procedural causality in episodic recollection. In Michaelian, K., Debus, D. & Perrin, D. (Eds.) *New directions in the philosophy of memory*. Routledge

Piccinini, G. (2022). Situated neural representations: Solving the problems of content. *Frontiers in Neurorobotics,* 16, 846979.

Poo, M. M., Pignatelli, M., Ryan, T. J., Tonegawa, S., Bonhoeffer, T., Martin, K. C., ... & Stevens, C. (2016). What is memory? The present state of the engram. *BMC biology,* 14, 1-18.

Pylyshyn, Z. W. (2007). *Things and places: How the mind connects with the world*. MIT press.

Quilty-Dunn, J., Porot, N., & Mandelbaum, E. (2023). The best game in town: The reemergence of the language-of-thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences*, 46, e261.

Radvansky, G. A., & Zacks, J. M. (2014). *Event cognition*. Oxford University Press.

Radvansky, G. A., & Zacks, J. M. (2017). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences*, 17, 133–140.

Ramirez, S., Liu, X., MacDonald, C. J., Moffa, A., Zhou, J., Redondo, R. L., & Tonegawa, S. (2015). Activating positive memory engrams suppresses depression-like behaviour. *Nature,* 522(7556), 335-339.

Ramsey, W. M. (2023). The Hard Problem of Content is Neither. *Review of Philosophy and Psychology*, 1-22.

Ranganath, C. (2022). Episodic Memory. In Kahana, M. & Wagner, A.D. (Eds.), *Handbook of Human Memory: Foundations and Applications.* Oxford University Press.

Rashid AJ, Yan C, Mercaldo V, Hsiang H-LL, Park S, Cole CJ, et al. Competition between engrams influences fear memory formation and recall. *Science* 353:383–387.

Recanati, F. (2012). *Mental files.* Oxford University Press.

Recanati, F. (2016). *Mental files in flux.* Oxford University Press.

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(7518), 426-430.

Refaeli, R., Kreisel, T., Groysman, M., Adamsky, A., & Goshen, I. (2023). Engram stability and maturation during systems consolidation. *Current Biology*, 33(18), 3942-3950.

Robins, S. (2016a). Optogenetics and the mechanism of false memory. *Synthese*, 193, 1561-1583.

Robins, S. (2016b). Misremembering. *Philosophical Psychology*, 29(3), 432-447.

Robins, S. (2016c). Representing the past: Memory traces and the causal theory of memory. *Philosophical Studies*, 173, 2993-3013.

Robins, S. (2020). Stable engrams and neural dynamics. *Philosophy of Science*, 87(5), 1130-1139.

Robins, S. (2023). The 21st century engram. *Wiley Interdisciplinary Reviews: Cognitive Science*, e1653.

Rolls, E. T. (2013). The mechanisms for pattern completion and pattern separation in the hippocampus. *Frontiers in systems neuroscience*, 7, 74.

Ross, T. W., & Easton, A. (2022). The hippocampal horizon: Constructing and segmenting experience for episodic memory. *Neuroscience & Biobehavioral Reviews*, 132, 181-196.

Rubin, D., Schrauf, R., & Greenberg, D. (2004). Stability in autobiographical memories. *Memory*, *12*(6), 715-721.

Rubinstein, D. Y., Weidemann, C. T., Sperling, M. R., & Kahana, M. J. (2023). Direct brain recordings suggest a causal subsequent-memory effect. *Cerebral Cortex,* 33(11), 6891-6901.

Ryan, T. J., Roy, D. S., Pignatelli, M., Arons, A., & Tonegawa, S. (2015). Engram cells retain memory under retrograde amnesia. *Science*, 348(6238), 1007-1013.

Sant'Anna, A., Michaelian, K., & Andonovski, N. (2023). Autonoesis and episodicity: Perspectives from philosophy of memory. *Wiley Interdisciplinary Reviews: Cognitive Science,* e1665.

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The future of memory: remembering, imagining, and the brain. *Neuron,* 76(4), 677-694.

Schacter, D. L., Eich, J. E., & Tulving, E. (1978). Richard Semon's theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 17(6), 721-743.

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind.* Cambridge University Press.

Semon, R. W. (1921). *The mneme.* G. Allen & Unwin Limited.

Shanton, K. (2011). Memory, knowledge and epistemic competence. *Review of Philosophy and Psychology*, 2, 89-104.

Siegel, G. (2022). Perceptual modes of presentation as object files. *Erkenntnis*, 1-19.

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

Sugar, J., & Moser, M. B. (2019). Episodic memory: Neuronal codes for what, where, and when. *Hippocampus*, 29(12), 1190-1205.

Sun, C., Yang, W., Martin, J., & Tonegawa, S. (2020). Hippocampal neurons represent events as transferable units of experience. *Nature neuroscience*, 23(5), 651-663.

Sutton, J. (1998). *Philosophy and memory traces: Descartes to connectionism*. Cambridge University Press

Sutton, J., & O'Brien, G. (2022). Distributed traces and the causal theory of constructive memory. In *Current controversies in philosophy of memory* (pp. 82-104). Routledge.

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology*: General, 138(2), 236–257.

Tanaka, K. Z., He, H., Tomar, A., Niisato, K., Huang, A. J., & McHugh, T. J. (2018). The hippocampal engram maps experience but not place. *Science*, 361(6400), 392-397.

Teyler, T. J., & DiScenna, P. (1985). The role of hippocampus in memory: a hypothesis. *Neuroscience & Biobehavioral Reviews*, *9*(3), 377-389.

Teyler, T. J. & DiScenna, P. (1986). The hippocampal memory indexing theory. *Behavioral Neuroscience*, 100: 147–152

Teyler, T. J., & Rudy, J. W. (2007). The hippocampal indexing theory and episodic memory: updating the index. *Hippocampus*, 17(12), 1158-1169.

Thompson, R. F. (2005). In search of memory traces. *Annual Review of Psychology*, 56, 1-23.

Tonegawa, S., Liu, X., Ramirez, S., & Redondo, R. (2015). Memory engram cells have come of age. *Neuron*, 87(5), 918-931.

Tonegawa, S., Morrissey, M. D., & Kitamura, T. (2018). The role of engram cells in the systems consolidation of memory. *Nature Reviews Neuroscience*, 19(8), 485-498.

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. (1995). Organization of memory: Quo vadis? In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 839–853). The MIT Press.

Tulving, E. (2002). Does Memory Encoding Exist?. In Naveh-Benjamin, M., Moscovitch, M., Roediger, H. L., & III (Eds.) P*erspectives on Human Memory and Cognitive Aging*. Psychology Press.

Tulving, E., & Watkins, M. J. (1975). Structure of memory traces. *Psychological review*, 82(4), 261.

Van Gelder, T. (1991). What is the "D" in "PDP"? A survey of the concept of distribution. In W. Ramsey, S. P. Stich, & D. E. Rumelhart (Eds.), *Philosophy and connectionist theory* (pp. 33-59). Hillsdale, N J: Erlbaum

Wang, Y. C., Adcock, R. A., & Egner, T. (2023). Toward an integrative account of internal and external determinants of event segmentation. *Psychonomic Bulletin & Review,* 1-23.

Wang, Y. C., & Egner, T. (2022). Switching task sets creates event boundaries in memory. *Cognition*, 221, 104992

Weidemann, C. T., & Kahana, M. J. (2021). Neural measures of subsequent memory reflect endogenous variability in cognitive function. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(4), 641.

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.

Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society*, 17(5), 766-780.

Xue, G., Dong, Q., Chen, C., Lu, Z., Mumford, J. A., and Poldrack, R. A. (2010). Greater neural pattern similarity across repetitions is associated with better memory. *Science* 330(6000):97-101.

Yates, T. S., Sherman, B. E., & Yousif, S. R. (2023). More than a moment: What does it mean to call something an 'event'?. *Psychonomic Bulletin & Review*, 1-16.

Yousif, S. R., & Scholl, B. J. (2019). The one-is-more illusion: Sets of discrete objects appear less extended than equivalent continuous entities in both space and time. *Cognition,* 185, 121–130.

Zacks, J. M. (2020). Event perception and memory. *Annual Review of Psychology,* 71, 165-191.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, 127(1), 3.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, 133(2), 273.

Zeithamova, D., Gelman, B. D., Frank, L., & Preston, A. R. (2018). Abstract representation of prospective reward in the hippocampus. *Journal of Neuroscience,* 38(47), 10093–10101

Zemach, E. M. (1983). Memory: What it is, and what it cannot possibly be. *Philosophy and Phenomenological Research*, 44(1), 31-44

Zheng, J., Schjetnan, A. G. P., Yebra, M., Gomes, B. A., Mosher, C. P., Kalia, S. K., Valiante, T. A., Mamelak, A. N., Kreiman, G., & Rutishauser, U. (2022). Neurons detect cognitive boundaries to structure episodic memories in humans. *Nature Neuroscience,* 25(3), Article 3.

Zhou, Y., Zhu, H., Liu, Z., Chen, X., Su, X., Ma, C., ... & Ma, L. (2019). A ventral CA1 to nucleus accumbens core engram circuit mediates conditioned place preference for cocaine. *Nature neuroscience*, 22(12), 1986-1999.