Semanticization Challenges the Episodic/Semantic Distinction

Sara Aronowitz

This is a penultimate draft of a paper forthcoming in *The British Journal for the Philosophy of Science.*

Abstract

Episodic and semantic memory are often taken to be fundamentally different mental systems, and contemporary philosophers often pursue research questions about episodic memory, in particular, in isolation from semantic memory. This paper challenges that assumption, and puts pressure on philosophical approaches to memory that break off episodic memory as its own standalone topic. I present and systematize psychological and neuroscientific theories of semanticization, the thesis that memory content tends to drift from episodic to semantic in structure over time and exposure to an environment. Semanticization, I argue, is a long-term interconnection between episodic and semantic systems that requires approaching both the content and function of these two memory systems as a whole. Thus we have a reason to reject projects by Michael Martin, which aims to carve out a uniquely episodic memory content, and Kourken Michaelian, which pairs episodic memory to its own unique function. Instead, seeing declarative memory as a single system with two facets or even a continuum of features allows for deeper insight into both content and function.

Introduction

Imagine[1] that yesterday, you went to a baseball game for the very first time. Today, you might remember where you parked, how you felt at the security station, and the crab fries you ate. But now imagine you took a job at that stadium, and worked there five days a week for a year. Would you remember where you sat yesterday? Some cognitive scientists have theorized that as you gain exposure to an environment, you have a tendency to shift from remembering it in first-personal episodes, to remembering it in an impersonal, factual manner. A shift from episodic to semantic modes of remembering, in other words. This phenonomenon is called semanticization. It may not be an absolute tendency, and could even turn out to be illusory – or a shallow phenomenon that does not reflect what's really going on in your head. But substantial evidence, as I'll demonstrate, suggests that we should take this idea seriously.

If so, I argue, there is a problem based in semanticization that challenges many recent philosophical approaches to memory. Namely, approaches that treat episodic memory as a distinct unit of philosophical study. I'll raise two versions of the challenge: one for stand-alone theories of episodic content, and the second for stand-alone theories of episodic function. In both cases, attending to the way semantic and episodic systems work together (if indeed they are

[1]I am grateful to Sarah Robins and Si-Won Song for incisive comments on this project, and to Thad Polk, whose class first got me thinking about this question.

even separate systems at all) opens up important questions that are invisible on the stand-alone episodic picture.

I'll set the stage in § 1 by surveying some of the ways the semantic/episodic divide has featured in philosophy, and dismissing a preliminary objection that episodic and semantic memory are clearly distinct natural kinds. Then in § 2, I introduce the phenomenon of semanticization and discuss its various forms and connection with larger theories of cognition. My negative argument has two parts: in § 3, a challenge to Michael Martin's view of episodic memory as a distinctive sort of content, and in § 4, a challenge to Kourken Michaelian's view of episodic memory as having a distinctive sort of function.

1 Dissociation

The contemporary semantic/episodic distinction in memory research was originated by Endel Tulving (1972), and was quickly assimilated into work on memory in psychology and elsewhere. Episodic memory, in its most paradigmatic form, is a vivid bringing to mind of a past incident in quasi-experiential format – almost a mental movie of something that happened to you. It can be thought of as indexed to the self in two ways: in that it is organized around a perspective, and is accompanied by an 'autonoetic' feeling, a sense of reliving the past. Both of these components can occur outside of memory as well. The first, in perception and some kinds of imagination, as we experience the world in a perspectival way, from our own location in space and time, and from our own body. The second in deja vu, where we have a non-memory experience that still has the feeling of past-ness. This distinction may also follow a more common-sense one, where in English, for instance, we would not usually say 'Leidy remembers the fireworks at the fair' if Leidy had purely semantic knowledge about a fair that she had never in any way experienced. Instead we would likely say "Leidy knows there were fireworks at the fair".

In this paper, I will not aim to give a comprehensive definition of the episodic/semantic distinction. While most, if not all, memory researchers would agree the cases described above are paradigmatically episodic and semantic, there is far less agreement about which features are most essential for making the distinction, and thereby how to handle cases where all the features do not coincide. Difficult cases include vividly perceptual memories that are untethered from a past episode, autobiographical memories that are fact-like rather than experiential, and memories of objects or people rather than events. This difficulty is compounded by a debate about episodic and semantic memories in non-human animals. On the one hand, Mahr & Csibra (2018) argue that non-human animals likely do not possess episodic memory at all due to an absence of autonoesis ("Thus, from our perspective, it seems unlikely that other animals (and very young children) would have the capacity for entertaining autonoetic memories, simply because they do not need it."). On the other hand, most of what we know about the neural structures of human episodic memory comes from work on the rodent medial temporal lobe, and many other researchers hold that rodents and other animals do have episodic memory based on anatomical similarities (Eacott & Easton, 2010) as well as their ability to pick out and learn from individual past episodes, rather than statistical or other forms of general information (Gershman & Daw, 2017). Fortunately, semanticization can be defined without adjucating these debates. As long as we agree on the paradigm cases, it should be possible to operationalize moving between these paradigmatic poles over time without determining where the point of transition from one type to another should lie, and even which features are ultimately most significant or essential for the categories.

Many (including Tulving himself) do not see the episodic/semantic distinction as carving up two fully separate systems so much as interacting parts of the same system, and others, such as Rubin (2021) have offered alternative ways to think of the relationship between the many

features of "episodic" and "semantic" memories. But often the distinction has become reified in a way that justifies a methodology of independent research on the two types of memory. Stanley Klein (2014) has gone so far as to suggest that not only are these distinct kinds of thinking, but that only episodic memory is genuinely memory at all. More common is the idea that while episodic and semantic denote two kinds of memory, we can ask philosophical questions of the two systems in isolation. In practice, this mostly entails focusing on episodic memory. Perhaps for some of the reasons Klein highlights, the implicit assumption seems to be that episodic memory is somehow more philosophically interesting: it has a unique phenomenology, it has more of a chance of being localized to a single neural or cognitive system, and it raises a restricted and unique range of epistemological issues. Michaelian (2016), for instance, initially justifies his focus on episodic memory by noting two considerations: (a) prima facie differences in content and phenomenology between the two systems, and (b) the potential to use episodic memory to explore the underexplored area of non-propositional knowledge. He concludes: "Thus it seems advisable to adopt a piecemeal strategy, initially developing separate accounts of episodic and semantic memory and reserving the task of combining them for future research".

Before discussing semanticization, we need to address a background issue: if there's a robust distinction between episodic and semantic memory biologically, then doesn't that answer our philosophical question? In other words, if semantic and episodic memory were separate natural kinds, shouldn't that mean scepticism about their separation is unwarranted? I will not say anything here about the conditional either way, but I want to head off the idea that we have a clear biological basis for the distinction.

A classic test of distinct neuropsychological function is double-dissociation: consider two functions A and B, do we observe animals (intervened-on or naturally) who have a deficiency in A but B is unimpaired, and other animals who have a deficiency in B but A is unimpaired? This test in principle is fully behavioral, though in practice it is used in attempts to identify distinct neural systems. A more direct neural measures might be, for instance, localizing functions to separate brain regions, though this test has the problem that neurally connected functions may cross brain regions (e.g. large scale networks), and conversely the same region may subsume two very distinct functions.

Some researchers have posited a single or double dissociation in the case of the episodic/semantic distinction (Smith et al 2011; Temple et al 2004). Typically, the two impairments are semantic dementia (SD), a form or syndrome of fronto-temporal dementia, and medial temporal lobe (MTL) amnesia, a memory impairment typically observed following injury, stroke, or surgical lesion (though in addition to SD, some researchers have focused on brain injuries and epileptic damage that might cause specifically semantic impairments). The putative dissociation is that semantic dementia preserves episodic memory while semantic memory is degraded, whereas MTL amnesia preserves semantic memory while episodic memory is degraded.

However, this story does not fit aspects of what we know about these impairments. It seems that most or all SD patients have episodic impairments, and perhaps more significantly, there is evidence that these impairments scale with their semantic impairments (Maguire, Kumaaran, Hassabis, & Kopelman, 2010)). This correlation is especially important given that most (arguably, all) of these patients have some semantic function remaining, so it would not make sense for dissociation to require a complete incapacity to remember episodically. On the other side, MTL amnesics have deep semantic impairments, except in cases when their episodic memory impairment is not that severe (Squire & Zola, 1998; Greenberg, Keane, Ryan, & Verfaellie, 2009).

These data are sufficient to show that we do not have clear evidence of a double-dissociation, or even a single dissociation. But they are not so surprising, once we look closely at the battery of tests used to assess each impairment. Suppose I asked you to tell me about your high school

graduation. You might just instantly remember a clear episodic scene of that day. But you might also think about general features of graduations, and use them to elaborate an answer: did you wear a cap? Was your grandmother there? Would this have been inside or outside, based on the average weather in Philadelphia in June? And so on. Likewise, were I to ask you why calico cats are always female, you might think back to a time when your mom explained the answer, or remember sitting in a middle school classroom, or even think back to your own encounters with calico cats to verify that it's true that all are female. That is, you use semantic and episodic strategies to answer both "semantic" and "episodic" questions, often in combination with one another. This does not necessarily mean that episodic and semantic memory are deeply related, but it does point to a difficulty in probing their connection, and the inadequacy of some of our current methods.

In sum. Double dissociation is a classic test of a natural division between neurocognitive systems. In the case of the semantic and episodic systems, a true double dissociation has been illusive. This does not mean that there is a single united declarative memory system, but it does entail that we do not have a strong scientific basis right now to accept the episodic/semantic distinction as one between products of two distinct natural kinds of cognition. For our purposes, this is not the end of the question about the relationship between episodic and semantic, but the beginning.

2 Semanticization: what and why

Semanticization [2] is the name for both a particular mental or neural occurrence (a memory shifting from semantic to episodic) and a general hypothesis about memory (that content in general tends to shift in this direction over time and exposure). While it is hard to deny the occasional occurrence of semanticization, the harder and more interesting question is about the general trend. One question, of course, is just whether such a general trend exists. If it does, we can then ask: is this shift a transfer of information within memory, or does it occur through some other route?

On the question of a trend, two main lines of evidence have been put forward. First, individual studies in particular memory domains, and second, broader theories about the rational function and evolutionary development of memory systems. Within each line of evidence we might articulate the semanticization hypothesis at different levels. For instance, that there is a transfer of information between brain regions associated with episodic tasks to those associated with semantic tasks. Or instead, that there is a shift in behavior around the same type of memory content, shifting from episodic-typical verbal descriptions to more semantic ones. These hypotheses are not unrelated, but we should be careful to distinguish them. I'll focus on two levels of analysis here:

neural semanticization information about past environments is initially encoded in neural systems associated with episodic storage and retrieval in the medial temporal lobe, but over time and exposure, this information shifts or is moved from the MTL and instead depends on neocortical areas associated with semantic storage and retrieval

cognitive semanticization information about past environments is initially stored and retrieved in an episodic form, with features such as concrete detail, imagery, and connection with other temporally related content, but over time and exposure, this information shifts away from these episodic features and is consolidated in an increasingly semantic

[2] see Andonovski (2020) for a different philosophical use of semanticization

> form characterized by gist, abstraction, and connection with other semantically related content.

Both definitions reference the idea of the very same information "moving" from one type of encoding to another. At a minimum, this means the information is initially present in one location or format and then subsequently in another location or format, but most researchers working on semanticization seem to have in mind something stronger: that there is a mechanism causing the "move", rather than merely an appearance and disappearance.

I'll start with direct evidence of semanticization in particular domains, and then survey two big-picture theories of memory that entail semanticization at both the neural and cognitive level. Semanticization at the cognitive level was already documented by Bartlett (1932). In the "War of the Ghosts", he describes some of these transformations anecdotally as following this trend: "Each illustration so far given shows a tendency to abbreviate and simplify both the story as a whole and also all the details that are reported."(72). This tendency goes along with an attempt to make sense of the story through rationalization and conventionalization. Thus we see a behavioral drift from specific, concrete, and experiential memory to a more gist-based, streamlined, and conventional version. Bartlett is particularly interested in how these stories don't just reduce in detail and resolution, but shift toward forms that reflect general knowledge about the domain and schemas for stories themselves.

In our terms, if the story were to just get fuzzier and less detailed over the years, this would not indicate semanticization. If Bartlett was right, however, the trend is to move away from particulars of what was presented and to incorporate information from other episodes, and so is at least related to the cognitive semanticization hypothesis. Of course, Bartlett's methods depart from our contemporary standards, but these findings were also replicated in a more controlled (but shorter) design by Bergman & Roediger (1999), and in a related but distinct form by Roediger et al. (2014).

More recent work has pursued several kinds of cognitive tests for semanticization. One family of measures uses direct reports, where semanticization would entail a shift from describing the target domain in terms of "remembering" to describing the domain in terms of "knowing". Somner (2017) uses this approach alongside fMRI imaging, and finds exactly that trend over the course of a 302-day study, correlated with a shift from hippocampal to neocortical activation. Relatedly, the cognitive semanticization hypothesis predicts a shift from experiential, detail-oriented language to objective, general language in describing past experiences. Salasoo et al. (1985) document a linguistic phenomenon they call "codification" which is an instance of semanticization of linguistic meaning. Finally, reaction-time, errors, and other cognitive measures have been linked to semanticization as well; for instance, Raaijmakers (2004) provides a semanticization-based model that explains the "spacing effect": that memory for an item is typically enhanced when it is shown twice separated by a longer interval of time, rather than a shorter one. Reagh and Yassa (2014) focus on competition effects between similar memories. They show that participants' memory for an item changes as they see it more often; as they were shown an item repeatedly up to 3 times, the ability to recognize the item increased but also the likelihood of falsely identifying similar items. This finding is predicted by the cognitive semanticization hypothesis because semanticization should increase the salience of gist, and thereby increase the rate of false alarms on semantically-related items.

At the neural level, it is relatively uncontroversial that neocortical activity increases over time and exposure. However, neural semanticization also requires that hippocampal (or other MTL) activity decrease at the same time. This latter point is far more controversial, and also raises a further question: do fully semanticized memories ever become entirely independent of the MTL? Piolino et al (2004), using PET, actually found an increase in hippocampal activity for

more remote memories than for recent ones, whereas Nadel et al. (2007), using fMRI, found that hippocampal activity did not change significantly over the course of time or over repeated retrieval (though cortical activity did increase with more retrieval events). These findings can be made consistent with semanticization if we assume that the particular remote memories being triggered here are still episodic, and that due to the structure of the tasks we are not selecting randomly from memories from a certain time but biasing search towards more episodic memories. To get into the details of this debate, we'll thus need to move into a broader discussion of theories of memory function that draws on both the neural and the cognitive level. I'll discuss two approaches in this vein.

2.1 Complementary Learning Systems

The Complementary Learning Systems (CLS) hypothesis, first formulated by McClelland et al (1995), provides a theory of *why* memory changes over time, rather than just where these changes can be localized to. The details of the view have changed over time but the core idea is that we have two different structures for encoding memories, one in the hippocampus and a second in the neocortex. The hippocampal structure is optimized for encoding representations of events inherited more or less directly from experience. These representations are specific, detailed, and are quickly encoded. The neocortical structure is optimized for encoding what McClelland et al. call structure, information that crosses over between individual events and objects, and forms the basis for concepts, categories, and schemas. This system learns much more slowly, and the learning process is the source of much of the structure. The learning process itself involves the hippocampal system replaying episodes over and over to the neocortical system, to allow for the slow but deep absorption of the structural information.

This theory is supported (and can likewise be contested) by the kind of direct neural and cognitive evidence discussed above. But it also aims to explain why we should process information with these two systems. Using a comparison with neural networks, McClelland et al. argue that the CLS framework is an effective answer to a learning problem we face: we take in experience in a spatio-temporal, perspectival format, one event at a time, and yet, we must use this information to make inferences from properties of many events, and even properties that span event types and levels. With just a hippocampal-type fast and faithful system, we could perform inferences to deduce generalizations when needed, but this would be computationally complex and time-consuming. With just a neocortical-type slow and structural system, we would not have time to encode details of events that might be relevant later and would be overwhelmed by the flow of experience. When both systems work together through replay, we can consolidate information in a structural form off-line – that is, instead of doing a live, complex computation at the moment some topic is raised, we can somewhat passively unload the information from the hippocampal system during sleep and in the background during waking.

Why does semanticization occur on the CLS theory? Consider a person who has many encounters with a baseball stadium over time. The CLS theory predicts that first encoding is always episodic, through the MTL system. Two things happen over time. First, through replay of many episodes of exploring the stadium, features of the stadium become semantically linked with other conceptual knowledge – for instance, the opening hours of the food stand might be assimilated into a general pattern of stores in small towns being closed on Sundays. Second, within the representation of the environment itself, the same process might occur at a finer grade of granularity: as he starts to visit the stadium at different times and through different entrances, the projection of these episodes allows links to form between less spatio-temporally close observations of the same environment. We might consider these two changes to be the same thing at different levels of abstraction, though it's also possible to treat the later as a form

of spatio-temporal or event-based abstraction rather than the more straightforwardly semantic abstraction in the first. But either would be sufficient to predict increased semantic properties at retrieval in many cases.

But what about deterioration of episodic retrieval? Here I am not sure what the CLS should say. One response could be that increased semantic retrieval masks episodic retrieval just due to the competitive nature of retrieval processing. The more controversial elements of the CLS theory concern the involvement of the hippocampus over time in memory, which as I've noted has been the subject of conflicting neural evidence. It may be that CLS theory is best formulated as not only a teaching of episodic-type information to the neocortex but also the gradual loss or supression of that information from the hippocampus – but the debate is evolving quickly as we acquire new evidence and methods.

I've skipped over many of the interesting details here, but this theory remains viable though controversial (see O'Reilly et al. (2014) for an opinionated update).

2.2 The Navigational Theory

Another theory that aims not just to identify semanticization but to explain why it would happen is what I'll call the navigational theory. Buzsaki and Moser (2013) propose that we should understand human and other animal memory systems as originating in navigational systems, and as preserving the same fundamental computational features. In navigation, we can distinguish between two families of strategies: egocentric and allocentric. Egocentric strategies identify which way to go by reference to a position in space – for instance, "turn left at the gate, and the cabinet will be on your right". Allocentric strategies indentify a location based on features of the environment, rather than the position of the person – for instance, "the cabinet is in the northwest corner of the central hallway".

In spatial navigation for both rodents and humans, egocentric strategies generally precede allocentric ones, and the balance shifts toward allocentric over exposure to the environment (Ekstrom & Isham, 2017) – a pattern that has been replicated in artificial agents as well (Uria et al 2020). The navigational view of semanticization uses this shift as a model for a shift from episodic to semantic. This amounts to a view about computational similarities between egocentric/allocentric and episodic/semantic, as well as about a shared biological and even evolutionary basis for navigation and memory.

Why, on this view, does semanticization occur? In navigation, egocentric strategies specify paths through a space. Imagine we had a piece of paper with many different paths drawn on it; eventually this would start to resemble a map. In more or less the same way, Buzsaki & Moser suggest that a node that is crossed once will be tied to the directionality of the path, but once it is approached and crossed from many directions, it becomes challenging to separate any particular path, and conversely begins to take on path-independent relationships with other nodes. That is, in the navigational case, we might think of the shift as passive, a property of the task and representation type rather than a target of active effort. To the extent that the episodic/semantic distinction reflects a similar computation, we should expect semanticization.

In both the CLS framework and the navigational theory, episodic memory serves as a kind of gateway or triage system into semantic memory. Both predict semanticization, though the process is somehow more active in emphasis in CLS. The two frameworks may even be consistent with one another, though they emphasis different features of the episodic/semantic divide. More generally, fairly strong evidence has accumulated for semanticization at both the neural and cognitive level. However, this is far from enough to be sure that semanticization is a profound trend and not just an occasional occurence. Further, it seems that to fully understand what semanticization might be and why it occurs, we will need a more complete theory of the

episodic/semantic distinction itself. But for the purposes of this paper, I only aim to establish that our current evidence makes semanticization a live possibility.

3 First Challenge: content

The first challenge from semanticization goes against the idea that episodic memory can be treated as a distinct system for philosophical purposes since it has a distinctive type of content. While many philosophers merely assume that episodic memory is a distinct object of study, it would be more charitable to address work that justifies its focus on the episodic. For this, I'll take Michael Martin (2001) as my example, since I take him to give a compelling case for episodic memory as a standalone kind, and one that is in contact to some degree with a scientific conception of memory.

Martin argues that were we to accept the idea that episodic and semantic memory are actually a single kind (a view he calls "constructivism", though this term now has a different meaning in philosophy of memory), we would miss out on an distinctive feature of episodic memory. Rather than rejecting constructivism outright, he is careful to conclude only that there would be an explanatory gap if we did so.

On his view, semantic and episodic memory both provide knowledge of the past, but in different ways. When I remember that there is a city called Hyderabad in Pakistan and in India, I also know that there are these two cities. Semantic memory entails knowledge[3], and in fact is virtually synonymous with knowledge. Episodic memory, on the other, does not take a proposition as its content, but instead relates us to a past event. In remembering, we retain our apprehension with that past event, where apprehension is a relationship that holds between us and events, as well as objects and places. Apprehension differs from acquaintance in that it is essentially episodic, referring to a relation at a specific time rather than a generic connection of long spans of time. Episodic memory does not entail knowledge by acquaintance of its object, because, for instance, I may have known Hyderabad then without still knowing it now. Instead, episodic memory is a representation of past experience, which we use in the good cases to retain our acquaintance over time, whereas in the somewhat-less-good cases, we merely preserve our apprehension of how things were.

Thus the distinctive features of episodic memory, for Martin, all stem from its content: because its content is events rather than propositions, it relates us to the world in a temporally specific way and is tied to a non-propositional form of knowledge, i.e. acquaintance. Because this content is represented as a form of past experience, episodic memory relates us to particular things, rather than generic or underspecified ones. To see this last point, consider that it's possible to know (and semantically remember) that a cat is on the table without it being any particular cat, but if I were to apprehend, and later episodically remember, a cat on the table, this would imply a relation to a particular cat. Thus, for Martin, the features of non-propositional structure and particularity of episodic memory derive from its content type.

Semanticization implies a range of contents differing in degrees of episodicity, rather than a precise boundary. Is this already a problem for Martin? As long as a substantial amount of our memories are purely episodic, his account would remain intact, and to deal with the in-between cases, he could leave them out entirely or treat their epistemic contribution as mixed itself. So vagueness is not itself a challenge. But semanticization affects the content of all episodic memories in the following way: episodic content must be the kind of thing that can be semanticized. This might seem like a minimal constraint – how could some content not be semanticizable?

[3] a view that perhaps only philosophers take seriously

Here we'll need to dig in to the specifics of the mechanism of semanticization. The navigational view is what we can call a passive view of semanticization: it claims that whatever it is that explains why memories become semanticized is not a form of active cognitive processing, but instead a natural and in some sense passive consequence of the accumulation of information of that kind. This is important on the navigational view, because the passive nature of the process allows the shared structure of egocentric/episodic and allcentric/semantic content to *explain* the transformation.

Therefore, only a certain kind of content supports semanticization: content that is episodic in the structural sense, mirroring egocentric content. What kind of content is this? The answer depends on the details of the navigational view. Perhaps both kinds of initial encoding are fundamentally ordered in a way that favors one direction, or they are perspectival in the sense of having an experiential centering, or something else entirely. In any event, the navigational view must posit some such structural similarity.

This constraint means that Martin's view, if it is be consistent with the navigational theory of semanticization, must consider how episodic content can support the transition to semantic content. In navigation, both strategies have the same kind of content (ways to get to the destination) in different forms. In Martin's view, the two types of content (events, propositions) do not seem to fall under any more general category. Another point of tension is that episodic memory taking an event as opposed to a proposition as its object does not fit the metaphors of the navigational view very well: a path through a map as opposed to a map are not oriented toward different kinds of knowledge, and there is no construct to take the place of the "node" that switches its role from path-dependent to path-independent.

The CLS hypothesis is more active than the navigational view, in that, using McClelland et al's metaphor, the hippocampal system "teaches" the neocortical system. This view still constrains episodic content however: recall that the hippocampal system is specialized to encode faithful representations inherited from experience, whereas the neocortical structure is specialized for structured, abstracted content. This difference is at the heart of the CLS theory, because it explains the difference in underlying neural features between the two systems as well as the need for two systems at all. And yet it is not obviously consistent with Martin's division in contents between episodic and semantic: at the very least, it articulates a conceptually distinct systemization of the difference between episodic and semantic content.

To see the issue, let's consider a representation of the event of the cat resting on the table. This representation, according to the CLS theory, is episodic insofar as it is structured according to how it was experienced: an occurrence over time, related closely to other temporally and spatially near memories rather than to semantically or conceptually similar ones. For Martin, it is episodic insofar as it concerns my relation to this particular cat at a moment in time. Several differences can be seen between the two that make them, if not fully orthogonal, still somewhat distinct. The cat being this individual, Begemot, is crucial for Martin but seems irrelevant for the CLS view – if anything, identifying this cat as Begemot, the one I saw last week in the park, imposes a semantic-type distal relationship between spatially and temporally disparate memories. Conversely, that the memory be encoded in a way that contains the right kind of richness for feature extraction in replay is critical for the CLS view, but seems less important for Martin.

These difficulties may not be fatal. Instead, the lesson is that semanticization, on the navigational and CLS views. implies the kind of relation between episodic and semantic memory that directly and substantially predicts the type of content episodic memories have. As such, it is a mistake to break off the subject of episodic memory content and study it in isolation.

These difficulties point to something deeper amiss. Martin thinks of episodic and semantic memory as not just different in content, but as aiming at different epistemic ends. The end of

being acquainted with the world, and knowing propositions about the world, are related but distinct – it is possible to fully satisfy one without fully satisfying the other, and it may even be possible to be fully propositionally knowledgeable without being acquainted with anything at all. Semanticization, on the other hand, points to a dynamic relationship between the two systems in service of a joint goal. Over time and experience, we transform the way we think of our environment, and it is this transformation that fits awkwardly into a picture on which episodic and semantic memory contribute different elements to our knowledge. Knowledge by acquaintance and of propositions, then, may be truly distinct and irreducible, but it would be a mistake to conclude that this difference is supported by synchronic differences in episodic and semantic remembering: looking at longer-term changes, we see that these memories are likely the product of an interconnected system for keeping track of the world as we move through it. These two kinds of knowing, then, might be better understood as phases or elements of this joint system rather than lining up with phenomenal differences in its products. This leads us to the next section.

4 Second Challenge: function

In his book (2016), Kourken Michaelian provides a theory of episodic memory function that, much like Martin's, treats the problem of episodic memory function as a distinct question from that of semantic function. Again like Martin, Michaelian discusses his decision to do so, and justifies it based on a survey of empirical work on differences between the two types of memory. He qualifies the project of focusing on episodic memory specifically as a preliminary approach, conditional on future scientific understanding of the connection between semantic and episodic systems.

Michaelian's view of episodic function is defended in various ways across the book, and it's beyond the scope of this paper to discuss them all. His position is that episodic memory is a kind of imagining: a simulation of the past. The psychological and functional profile of episodic memory is continuous with that of imagination, and rather than preserving knowledge, episodic memory is a constructive faculty that pulls together all kinds of information to project what might have occurred. But if episodic memory is, like imagination, an activity I engage in while recollecting, how do I preserve information from the past? Or in other words, if episodic memory is merely an occurrent thought process, what faculty is responsible for holding on to knowledge over time?

Michaelian's answer: semantic memory. Or rather, semantic memory along with contextually available information at the time of recollecting are used together to create an imaginative simulation. He does at times seem to allow that episodic information can be stored, but denies that this storage is necessary for episodic remembering. By linking episodic memory and imagination, memory becomes an occurrent activity, only incidentally associated with storage and consolidation.

Semanticization raises an immediate challenge to Michaelian's view. Semanticization, of either the neural or cognitive variety, is a phenomenon that links episodic storage and retrieval to semantic storage and retrieval. Michaelian might at times allow for the possibility of episodic storage and retrieval, but he is adamant that this is not necessary for episodic remembering, and it does not fit in well with a picture of episodic memory as imagining. Semanticization is meant to apply to episodic memories in general – if it were an account of peripheral or special kinds of episodic memory, it would be so trivially satisfied as to be vacuous. So semanticization cannot describe a relationship between episodic and semantic memory at all on Michaelian's view. Episodic storage and retrieval is simply too peripheral to episodic memory to support such a connection.

Might Michaelian respond by arguing that semanticization can obtain even in occurrent episodes of thought? For instance, we could define semanticization as a shift between two types of episodes of remembering, such that the same subject-matter would be more likely to be remembered in an episodic mode (that is, constructed in a simulation) earlier and in a semantic mode (that is, directly retrieved or else inferred/constructed conceptually) later[4]. But this redefinition reveals two conceptual difficulties. First, on the face of things, semanticization is a change in how an event is remembered. But without anything like a trace or causal link on Michaelian's view connecting subsequent representations of the event to one another, it is hard to say what would link the later, semanticized memories to the earlier memory. At best, semanticization could be described as a change in the distribution of what is remembered about a past environment, rather than a transformation in how that environment is remembered. But this might be recast as a revision to our common-sense misconceptions about memory.

Second, for semanticization to be a phenomenon that holds over repeated exposure to the environment or over time, it must be a matter of dynamics *between* episodes, as well as *within* episodes. There needs, in other words, to be an effect of time and accumulation on remembering. However, any explanation for this effect would naturally concern what happens to information about an environment as it piles up, as time passes, and so on. But these are storage-based features. Episodic simulations, as far as Michaelian's discussion goes, do not pile up in memory stores, nor do they travel through long periods of time: instead they seem to come into existence when remembering, perhaps alter our thinking in some way, and then they are over. So even if we reconceptualize the long term effect of semanticization as over rememberings rather than stored memories, any explanation for the effect over time and accumulation needs to make contact with the information storage level. And while semantic memory does play this function for Michaelian, he says so little about it that it is unclear whether it could ground any such explanation.

A related problem is the breakdown of the parallel between episodic and semantic. The two frameworks discussed above, CLS and navigational theory, both treat episodic and semantic memory in parallel. Even supposing Michaelian were to fix the problem I've just noted by adapting a notion of semanticization to occurrent episodes of thought, he would then have a mismatch. On the one hand, episodic memory is a kind of occurrent thought, and on the other, semantic memory is still a more classical system for information storage and retrieval. Now he would have to address the mismatch on the semantic side: how could semanticization be a shift from one system to the other if these systems are not of the same basic kind? It would seem to be a category mistake to claim that content can shift from a way of imagining to a mechanism for storage. After all, as I noted already, ways of imagining (at least of the epistemically oriented variety) typically depend on mechanisms for storage already.

The sleight of hand in Michaelian's view is to make episodic memory into a fully simulatory system, the background work of carrying information through time must be outsourced to another cognitive system. Otherwise, we would be total amnesics. But it is this background, offline processing that characterizes the dynamics of semanticization. As such, it is at odds with Michaelian's theory.

5 A positive direction

I have argued that semanticization undermines an entrenched way of asking philosophical questions about memory. But beyond this negative implication, semanticization can also be the basis for a positive intervention in current debates around episodic confabulation. As diagnosed by

[4]thanks to an anonymous referee for suggesting this response on Michaelian's behalf

Robins (2016, 2019), the label 'confabulation' has been applied to a varied set of memory phenomena, from pathological faulty memory associated with psychiatric illness to benign cases of what she terms 'misremembering', where inaccurate (or merely unfaithful) details are added to a memory based on contextual cues.

A core component of these misremembering cases is that information from the environment is taken and encorporated into an earlier memory, resulting in a distortion. More specifically, if we look at the way information and memory are operationalized, we find that the information is typically semantic while the memory is episodic. For instance, Frenda et al. (2012) showed participants altered images and captions of real events, such as 'February 12, 1999: Speaking on the Senate floor at the conclusion of President Clinton's impeachment trial, Senator Joseph Lieberman, D-Conn, announces that he will vote guilty on the charge of perjury.'. Participants rated their agreement with statements such as: "I remember seeing this," and "I don't remember seeing it, but I remember it happening". Summing across several different events, over 20% of subjects reported remembering see the events. In this study, the misinformation came in a (partially) semantic form, an assertion about the event, and it would seem that remembering seeing would signal an episodic form of recollecting.

We might, then, call the process at work in this type of misremembering *episodicization*, meaning an inverse process of semanticization where semantic information is transformed into an episodic form. Episodicization would also encompass cases where a fact is used to create an episodic representation without any misinformation, such as when I lose track of what kind of dog I saw last week and later recollect it (correctly) as a black lab based on background knowledge of which species are most common.

Conceptualizing confabulation this way allows us to ask important new questions. There is much recent interest in why we make these kinds of memory errors: De Brigard (2014), for instance, arguing that we do so because of the future-directed nature of the memory system, and Robins (2016) holding that neither constructivism nor the more traditional archival conception of memory can make sense of misremembering. This debate has focused entirely on episodic memory, and answering the question of why the errors occur in a way that is internal to the functioning of episodic memory. But once we see key cases of misremembering as semantic to episodic transitions, it seems this focus is too narrow. Instead, we can ask: why does information sometimes move back to episodic form, and why does this occur less often than the reverse? And further, what is the optimal balance of semantic and episodic formats, and how can this vary based on environments, goals, and cognitive capacity? Likewise, on this episodicization framework, we can distinguish a new normative dimension, with transitions from semantic to episodic being seen as more or less accurate and faithful embeddings of the source information.

6 Conclusion

Episodic and semantic memory are increasingly treated as separate systems for the purpose of philosophical theories of memory. This paper demonstrated that this trend is on poor empirical grounding. We lack evidence that epistemic and semantic memory are separate natural kinds, according to the standard test of double dissociation. More significantly, a well-studied but still controversial set of ideas in memory research, semanticization, undermines the idea that either the content or the function of episodic memory can be studied in isolation. Semanticization, a dynamic process thought to link episodic and semantic systems over time, informs what kinds of content episodic memories can have by foregrounding a major role for those contents. In order for semanticization to work, episodic contents need to align in the right sort of way with semantic ones, though what this right sort of way is depends on the theory of semanticization. For the question of function, we've seen how one prominent theory of episodic function makes

semanticization illegible, a kind of category mistake. More generally, to the extent that the two broader frameworks about semanticization are correct, we should look for a theory of episodic function that does not conflict with episodic function in relation to semantic memory. Seeing these systems as linked also may help make progress on thorny debates about certain kinds of misremembering, which can now be understood as semanticization in reverse, or episodicization.

The bigger picture issue is this. Looking statically at our memories, we might conclude we have these two very different kinds that are neatly separated. Even if the divide is exceptionally crisp at any given time, this does not rule out a dynamic relation between the two kinds that might affect every single memory eventually. By attending to the dynamic process of semanticization, we are forced to attend to the possibility of a dynamic relationship between episodic and semantic memory. This stronger this relationship, the less sense it makes to consider episodic memory a system with separate contents or a separate function. Dynamic relationships can be almost completely obscured by taking a static snapshot, but still are fundamental to understanding the nature of a cognitive system.

References

[1] Nikola Andonovski, *Singularism about episodic memory*, Review of Philosophy and Psychology (2020), 1–31.

[2] Frederic Charles Bartlett, *Remembering: A study in experimental and social psychology*, Cambridge University Press, 1995.

[3] Erik T Bergman and Henry L Roediger, *Can bartlett's repeated reproduction experiments be replicated?*, Memory & cognition **27** (1999), no. 6, 937–947.

[4] György Buzsáki and Edvard I Moser, *Memory, navigation and theta rhythm in the hippocampal-entorhinal system*, Nature neuroscience **16** (2013), no. 2, 130–138.

[5] Felipe De Brigard, *Is memory for remembering? recollection as a form of episodic hypothetical thinking*, Synthese **191** (2014), no. 2, 155–185.

[6] Madeline J. Eacott and Alexander Easton, *Episodic memory in animals: Remembering which occasion*, Neuropsychologia **48** (2010), no. 8, 2273–2280, Animal Models of Amnesia.

[7] Arne D Ekstrom and Eve A Isham, *Human spatial navigation: representations across dimensions and scales*, Current opinion in behavioral sciences **17** (2017), 84–89.

[8] Daniel L Greenberg, Margaret M Keane, Lee Ryan, and Mieke Verfaellie, *Impaired category fluency in medial temporal lobe amnesia: The role of episodic memory*, Journal of Neuroscience **29** (2009), no. 35, 10900–10908.

[9] Stanley B Klein, *What memory is*, Wiley Interdisciplinary Reviews: Cognitive Science **6** (2015), no. 1, 1–38.

[10] Eleanor A Maguire, Dharshan Kumaran, Demis Hassabis, and Michael D Kopelman, *Autobiographical memory in semantic dementia: A longitudinal fmri study*, Neuropsychologia **48** (2010), no. 1, 123–136.

[11] Johannes B Mahr and Gergely Csibra, *Why do we remember? the communicative function of episodic memory*, Behavioral and brain sciences **41** (2018).

[12] Michael GF Martin, *Out of the past: Episodic recall as retained acquaintance*, Time and memory: Issues in philosophy and psychology (2001), 257–284.

[13] James L McClelland, Bruce L McNaughton, and Randall C O'Reilly, *Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory.*, Psychological review **102** (1995), no. 3, 419.

[14] Kourken Michaelian, *Mental time travel: Episodic memory and our knowledge of the personal past*, MIT Press, 2016.

[15] Lynn Nadel, Jenna Campbell, and Lee Ryan, *Autobiographical memory retrieval and hippocampal activation as a function of repetition and the passage of time*, Neural plasticity **2007** (2007).

[16] Randall C O'Reilly, Rajan Bhattacharyya, Michael D Howard, and Nicholas Ketz, *Complementary learning systems*, Cognitive science **38** (2014), no. 6, 1229–1248.

[17] Pascale Piolino, Gaëlle Giffard-Quillon, Béatrice Desgranges, Gaël Chételat, Jean-Claude Baron, and Francis Eustache, *Re-experiencing old memories via hippocampus: a pet study of autobiographical memory*, Neuroimage **22** (2004), no. 3, 1371–1383.

[18] Jeroen GW Raaijmakers, *Spacing and repetition effects in human memory: Application of the sam model*, Cognitive Science **27** (2003), no. 3, 431–452.

[19] Zachariah M Reagh and Michael A Yassa, *Repetition strengthens target recognition but impairs similar lure discrimination: evidence for trace competition*, Learning & Memory **21** (2014), no. 7, 342–346.

[20] Sarah K Robins, *Misremembering*, Philosophical Psychology **29** (2016), no. 3, 432–447.

[21] _____, *Confabulation and constructive memory*, Synthese **196** (2019), no. 6, 2135–2151.

[22] David C Rubin, *A conceptual space for episodic and semantic memory*, Memory & cognition (2021), 1–14.

[23] Aita Salasoo, Richard M Shiffrin, and Timothy C Feustel, *Building permanent memory codes: codification and repetition effects in word identification.*, Journal of Experimental Psychology: General **114** (1985), no. 1, 50.

[24] Mary Lou Smith and Suncica Lah, *One declarative memory system or two? the relationship between episodic and semantic memory in children with temporal lobe epilepsy.*, Neuropsychology **25** (2011), no. 5, 634.

[25] Tobias Sommer, *The emergence of knowledge and how it supports the memory for novel related information*, Cerebral Cortex **27** (2017), no. 3, 1906–1921.

[26] Larry R Squire and Stuart M Zola, *Episodic memory, semantic memory, and amnesia*, Hippocampus **8** (1998), no. 3, 205–211.

[27] Christine M Temple and Paul Richardson, *Developmental amnesia: A new pattern of dissociation with intact episodic memory*, Neuropsychologia **42** (2004), no. 6, 764–781.

[28] Endel Tulving, *12. episodic and semantic memory*, Organization of memory/Eds E. Tulving, W. Donaldson, NY: Academic Press (1972), 381–403.