RHV, 2025, No 28, 92-107 eISSN 0719-4242
DOI: https://doi.org/10.22370/rhv2025iss28.4848
Philosophy of memory / Articles

When remembering gets it wrong: Memory assemblages and the content of what is retained

Cuando el recuerdo se engaña: Los ensamblajes de la memoria y el contenido de lo que se retiene

Hilan Bensusan*

Associate Professor
University of Brasilia
hilantra@gmail.com
https://orcid.org/0000-0001-6686-9751

Abstract

Misremembering is a common feature of human experience. De Brigard has proposed a model in which human episodic memory functions as part of a broader cognitive system oriented toward hypothetical processing. As a result, remembering and misremembering serve a similar cognitive function. This model is contrasted with Bensusan's pan-mnemism —a view that understands memory phenomena as ubiquitous and maintains that any act of retrieval depends not only on what has been retained but also on subsequent additions that shape the content. What is retained is therefore always subject to the influence of new events and thoughts. Pan-mnemism conceives of misremembering not only as fulfilling the same cognitive function as remembering but also as involving genuine episodes of memory. After comparing these two approaches, the paper presents elements of an argument suggesting that memory traces —and, more broadly, any kind of causal ontological intermediaries— are dispensable for both characterizing the function of memory and explaining human episodic (mis)remembering.

Keywords: De Brigard, misremembering, pan-mnemism, memory traces, hypothetical thinking

Received: 29/03/2025. Final version: 04/07/2025
eISSN 0719-4242

Resumen

El recuerdo erróneo es una característica común de la experiencia humana. De Brigard ha propuesto un modelo en el que la memoria episódica humana funciona como parte de un sistema cognitivo más amplio orientado al procesamiento hipotético. En consecuencia, el recuerdo y el recuerdo erróneo tienen una funcción cogniva similar. Este modelo contrasta con la articulación del pan-mnemismo de Bensusan, un punto de vista que considera que los fenómenos de la memoria son ubicuos y postula que cualquier acto de recuperación depende no sólo de lo que se ha retenido, sino también de las adiciones posteriores que dan forma al contenido. Por tanto, lo que se conserva está siempre sujeto a la influencia de nuevos acontecimientos y pensamientos. El pan-mnemismo concibe el recuerdo erróneo como un episodio de memoria. Tras comparar estos dos enfoques, el artículo ofrece elementos de un argumento que sugiere que los rastros de memoria – y, más ampliamente, cualquier tipo de intermediarios ontológicos causales – son prescindibles tanto para caracterizar la función de la memoria como para explicar el (mal)recuerdo episódico humano.

Palabras clave: De Brigard, recuerdo erróneo, pan-mnemismo, rastros de memoria, pensamiento hipotético.

1. Introduction

Memory is, on the face of it, an odd phenomenon. It always requires more than one moment in time that can be arbitrarily separated from each other. Further, memory cannot be said to be fully in none of the two or more distinct times that constitute it. It necessarily involves two moments in time: a time when something (say, *x*) is stored, kept, or retained, and a time when something (say, *y*) is brought back, recuperated, or retrieved. Several issues concerning memory arise from this diachronicity, which also involves a duality between *x*, the retained element, and *y*, the retrieved one. In apparent contrast with many other cognitive abilities —such as perception, attention, or reasoning— which can be described in terms of an exercise involving a background and an object, and which may be portrayed as unfolding within a single moment in time (whether brief or extended), memory is split into two distinct moments that are ostensibly separated from each other. As a result, we can speak of the retained *x* and the retrieved *y* and ask what the relation between them is —whether, for instance, it makes sense to say that $x = y$ or that *x* is somehow contained in *y*.

Efforts to understand memory often attempt to reduce its diachronicity to an underlying mechanism that simply endures in a single moment. Under this assumption, it makes sense

to suppose that *x* could gradually be transformed into *y*. Whether or not the relation between *x* and *y* is one of transformation, the question arises as to how best to characterize it.

This paper considers discussions about misremembering and its relation to the workings of human (episodic) memory. De Brigard (2014) proposes a model according to which misremembering shows that memory is not cognitively appropriate because it aims at remembering —it needs not to remember to be cognitively good. In this model, memory is part of a larger cognitive system oriented toward hypothetical processing. In contrast, the *pan-mnemist* model (Bensusan 2024) is brought to the fore. The latter conceives of retrieval as an assemblage of what is retained and what has been added in the meantime. These memory assemblages make retrieval a function not only of retained archives but also of *addends* that orient the extraction of retained content. This second model addresses misremembering — and imagination— without postulating an underlying cognitive system that contrasts with what we ought to understand as memory. The two models conceive of the content of *x* and *y* in different ways. While De Brigard's model posits intermediary traces connecting the transformation of what is retained into the retrieved *y*, the memory assemblages model requires no intermediaries beyond the *addends*, which are something quite distinct from memory traces.

The remainder of the paper is divided into four parts. First, I present the main features of De Brigard's model with respect to misremembering. I then contrast and compare it with the memory assemblages model in the following two sections. A final section addresses the issue of memory traces and concludes that they are either too vague or unnecessary

2. Misremembering and reasoning

Discussions about memory often stumble in the apparently frequent cases of misremembering. Recent work attempts to deal with actual cases of memory lapses and mistakes as considered in Brainerd & Reyna (2005), but the very issue of how imagination looks similar to memory is a puzzle for the study of memory at least since Aristotle (1972). The reliability of human memory is at stake when we consider how it seems to involve the systematic deployment of imagination and how it often fails to bring back events as they have actually happened. Cases of misremembering suggest that the relation between the retained (*x*) and the retrieved (*y*) is something other than that of identity or inclusion. What is retrieved in memory is often to be corrected and almost always incomplete. No matter if performed on purpose, remembering frequently appears tainted with attempts to fill in gaps, presumptions,

careless reconstructions, and outright inventions. The problem is more severe when we focus not on semantic memory concerning theories, hypotheses, or knowledge about the world, but rather on experienced facts or events —human episodic memory is taken to be dramatically unreliable to adequately describe what had been witnessed in the past. It is common for most people to mix what they have witnessed and retained with what they imagine that would make sense. This has led to questions concerning what kind of imagination is involved in the relation between *x* and *y* (see, for instance, Langland-Hassan 2021). At any rate, it is easy to conclude that memory, taken as a human capacity, is defective and then downplay it in comparison to other, artificial devices that are supposedly able to perform the task of memory in a manner that is both more efficient and more trustworthy. To be sure, a video recorder appears to be a more reliable archive of what took place in a bank robbery than the report of a witness who was standing nearby. The recording seems to engage neither in an effort to understand the past episode beyond what is archived nor in an attempt to make what matters for the subsequent occasion stand out. Its task can be circumscribed and arguably limited, but as a memory device dedicated to retaining something considered on its own, it would not misremember. The straightforward conclusion could then seem to be that human memory, compared to other devices that would be employed for similar tasks, is faulty and prone to misperform.

Such a conclusion, however, has been resisted in the light of both more detailed consideration of human memory and a more realistic picture of how devices like a video recorder are actually employed as an archive. Further, the idea that there are (or could be) memory devices that would be oblivious to the circumstances where retrieval takes place flies on the face of the idea that memory is a way to rely on what has been stored to guide action in the future (see Matthen 2010; Langland-Hassan 2021). A museum, a book, or a film (or a video file from a recorder) can inform their future audiences only in relation to what else has been stored since. Similarly, human (episodic) memory cannot be useful but in the light of other events stored somewhere else. If you're travelling in the desert and see an oasis in a location that, unbeknownst to you, has lightning that is apt to produce mirages, the perceived fact you have registered is justifiedly overruled. Human memory would be useless if it could not reconsider what is of importance given the circumstances of retrieval (Vosgerau 2010), and mechanical devices constitute memory only to the extent that they are properly documented and interpreted (Clark & Chalmers 1998; Clark 2008; Heersmink & Carter 2020).

It remains to be explained why human memory seems so unreliable as part of a more general problem concerning the relation between the retained *x* and the retrieved *y*. Misre-

membering comes in several different forms, such as boundary extension mistakes, confusion of relevant details, order mismatches, and the telescope effect, where the recent and the remote events get mixed together (see Neter & Waksberg 1964; De Brigard 2014). They indicate that (episodic) memory is something different from mere preservation of a content that encodes past events —and retrieval circumstances form a relevant part of what retrieving is (see Tulving 1983). As a consequence, the problem of misremembering paves the way to a more general investigation concerning the relation between retention and retrieval —and between *x* and *y*— once the diachronic activity of memory is clearly understood as being something other than that of preserving, that is, keeping *x* in the form of *y*.

To deal with this issue, Felipe De Brigard (2014) has proposed a model of human episodic memory as a component of a larger system involving episodic future reasoning and episodic counterfactual thinking. Such a system considers not only what has actually happened but also what previous experience reveals about what could have happened and what could take place in the future. De Brigard builds on the idea that, contrary to appearances, memories need not be faithful renditions of past events in order to be useful —they don't need to be past-preserving to be good. According to his model, memory is integrated into a modal system that enables the cognizer to reason hypothetically about both the future and the past. This helps explain the usefulness of remembering for future planning and prediction. De Brigard connects this capacity to an evolutionary account that privileges less accurate retention in favor of broader, modal reasoning about the environment (Sutton 2009). The model postulates that there are advantages to embedding memory within a broader reasoning system, such that storing events in a gist-like manner yields two kinds of benefits (De Brigard 2014, p. 166). First, it allows for better engagement with an environment that is constantly changing and rich in information —it makes sense, then, to have a retained *x* stored in a flexible or plastic format that can adapt to the changes that occur by the time of retrieval. Second, it is less costly in terms of storage to retain *x* in a form that is less than a mirror representation of the original event. While the first benefit concerns what memory must do in an environment where things are rarely stable, the second highlights the efficiency of a cognitive system that favors schematic encoding. De Brigard argues that memory is therefore part of a system for tracking information used in reasoning, for which remembering is not ultimately the main goal —memory, he claims, is not primarily for remembering. Drawing on Alva Noë's (2004) account of perception as action-guided, De Brigard maintains that memory should be understood in terms of its role in decision-making and action orientation. The broader system guides action by considering what would best fill the gaps in what has been retained, in order to reason about past events and assess the plausibility of future ones.

De Brigard's model explains why misremembering is not a cognitive failure but rather has to be considered in the light of what action in an unstable environment requires. It is for this larger purpose of a system geared towards action that memory produces adequate, or optimal, representations of events. De Brigard holds that given our need for action in an environment that lacks stability and the costs of keeping past episodes fully represented, "we cannot afford having every aspect of the world under the spotlight of attention" and thus "the next best solution is to have a system that can fill in the gaps with the optimal alternative it can come up with" (De Brigard 2014, p. 172). He understands that the memory we have is the second best alternative given that a content-preserving mechanism would be too costly and the environment seems to favor rough representations. Of the two benefits mentioned above —dealing with an unstable environment and using a less costly process of encoding— De Brigard seems thus to privilege the second while attending to the first merely as a consequence of the evolution of memory which arguably had taken place in an unstable environment. It follows that episodic human memory is not a geared towards memory; it is a different system altogether that could only be imperfect if compared to a system whose proper function is to optimize remembering. Hence, although De Brigard's model goes a long way to rescue memory from the bad press misremembering gives it, it fails to make justice to the idea that human memory is not deficient but often performs its proper mnemonic task.

The merit of De Brigard's model is to make explicit how memory and imagination are connected as both are arguably part of a larger system for episodic counterfactual thinking and episodic future reasoning. Imagining things mobilize a capacity to consider future and counterfactual episodes in a sense that is similar to what memory does. Both deal in judgments of plausibility and engage with modal considerations —both support counterfactual and future reasoning. The underlying picture of memory in the model, however, is that (episodic human) memory performs something other than what could be expected from memory —and perhaps imagination performs something other than imagination if we think of it in a similar vein as an organ for future reasoning and counterfactual thinking. In fact, it emphasizing that we have resorted to a second best alternative once we could not have, given the cost and circumstances, a proper memory doing what memory is ultimately supposed to do. As a consequence, it seems to assume as possible for a memory system to have a relation between the retained *x* and the retrieved *y* that is closer to an identity; we would then fail to have such a (better) system because of its costs and because of the inconvenience of having such a capacity in the unstable environment where our cognitive capacities hav evolved. That possibility, not implemented in human episodic memory for reasons that have therefore to do with cost-effectivity, appears as, in principle, a possible best option to which actual human memory is the second best. In contrast, I'll propose next an alternative account of memory that entails that no proper memory system can attain a retrieved *y* that is (close) identical to

the retained *x*. A memory system will appear as constituting a different relation between *x* and *y*, one that is responsive to the inevitable addends that are introduced since the (different) moment when retention took place. Once this is assumed, a different account of misremembering which will keep some of the important features brought up by De Brigard will be sketched. Human episodic memory will then emerge as a relevant example of a memory activity whose function is to do what should be reasonably expected from a memory system.

3. From a pan-mnemist point of view

Pan-mnemism is a view according to which memory is quite widespread in the world (Bensusan 2024). Its diachronicity, through which the two poles of retention and retrieval take place in two different moments, is not a peculiar event in a world of synchronous phenomena. It is, rather, a point of departure on which to model other states and abilities —including other cognitive states and other mental activities. Thus genetic codes, written text, geological and archaeological remains, the size of tree trunks as much as theories, reasonings or arguments, perceived landscapes and belief systems are understood in terms of two diachronic poles, one in which something is retained —the DNA of cells, the marks of the ink on paper or of code in a reading device, the first formulation of an account of something, the (arguably non-monotonic) inference from existing premises— and another one in which retrieval takes place —a sequence of proteins, a reading of the text, the reformulations of a theory in light of new evidence, the inference to a different conclusion given further premises. Pan-mnemism holds that these two diachronic poles are not exceptional but a basic component of what there is. Further, ubiquitous memory phenomena are understood as formed by what is retained and what is retrieved, which are separated by addends that have transformed the circumstances in which storage took place. These addends, which can be located between retention and retrieval, can take the form of both new information and new events that greatly reshape what has been retained. Because of the inevitability of addends, nothing can be preserved on its own. Also, it would make no sense to store anything if it is not for the benefit of something else —e.g., a different circumstance, the readers of the book, the visitors to the museum, or the future generation of cells.

Memory and addends are thus interconnected. No retention takes place on its own, and nothing can be retrieved once and for all —addends are needed both for retention and for subsequent episodes of retrieval. Whatever is retained in an organism, in a landscape, in a brain configuration is hostage to the addends to come; retention is thus somehow directed towards retrieval, which, in its turn, can act only on what has been retained. Retrieval is

performed through addends that never cease coming; what is retained can only be recuperated by them —a new reader, a novel individual, a different circumstance. A pair constituted of a collection of retained materials —called *archive*— and a retrieval episode —made by a collection of addends— form a *memory assemblage* which is then retained and beholden to future addends that will retrieve it in a different memory assemblage. These ubiquitous assemblages are modelled on the working of memory, which is always intertwined with the insistent arrival of addends. Memory appears as a basic component of what there is, and it relies on a diachronicity for retrieval always refers to a moment of the past. It follows that what pan-mnemism is capable of doing is to explain things in terms of one or more moments in time. It introduces a temporal duplicity (or multiplicity) into the basic understanding of things and situates human episodic memory in a landscape of memory assemblages. Time —which is arguably a consequence of the insistent addends— is a crucial component of these assemblages as it is argued that there is no memory without time intervals.

What can then be said about the relation between the retained x and the retrieved y? If memory and addends are interconnected, there is no easy way to formulate the relation between x and y in terms akin to identity. We can conjecture that y is a function of x and a collection of addends, but the very idea of a memory assemblage indicates that what is retrieved depends on more than a single retained element. What addends do is to reassemble what had been retained in different archives and then retrieve them in their own light. We can then formulate a memory assemblage as follows:

$$y = f(x_1, \ldots, x_n, a_1, \ldots, a_n),$$

where each a is an addend. The memory assemblage is then retained for further retrieval through different assemblages:

$$y' = f(y, x_1, \ldots, x_n, a_1, \ldots, a_n).$$

The relation between the retained x and the retrieved y is thus through addends and the assemblages they promote. The reading of a text, as the epigenetic conditions under which a genome is decoded, the bacterial environment in the trunk of a tree, and the new information added to a set of premises are addends that retrieve what they find retained by assembling a

collection $x_1, \ldots, x_n$ of archives. (It is clear that archives are themselves memory assemblages, as what is retained cannot be available for assembling without a retrieving action.) Addends, and the assemblages of archives that they bring about, are crucial to the relation between x and y —that relation is relative to what makes the diachrony between the two elements so that a retained x_i can be part of an indefinite number of memory assemblages and thus can be retrieved in an indefinite number of circumstances. The same retained element —the same geological site, the same remembrance, the same genetic code— could give rise to several retrieved assemblages. Thus, we seem to be very far away from the idea that the relation between x and y is akin to identity.

Memory assemblages bring together archives and addends that can be separated by indefinite spans of time. An archive can be mobilized in an assemblage at any time after retention; archaeological retrieval, for example, is carried out in light of what has been added since, and in conjunction with other relevant archives. When we consider the history of excavations at sites like Tiwanaku, near La Paz, Bolivia, we must take into account the new procedures that have been introduced —such as reliable radiocarbon dating methods— and the novel information about the Wari culture of Tiwanaku that has been acquired. Further, we must attend to the degradation of the site due to numerous physical and biological events that have occurred. Between the construction of the site —which is retained in the ruins— and contemporary archaeological inquiry, many things have been lost, including the population who preserved information about its foundation and maintenance. According to the pan-mnemist model of interaction between memory and addends, what has been lost is the effect of iterated addends (Bensusan 2024, pp. 168–73). Assembling memories occurs through the effort of investigation: one must rely on relevant archives and act within circumstances shaped by addends. Each account of the ruins forms a memory assemblage that dates the site differently. Around a century ago, the site was believed to have been built some 15,000 years ago, whereas today —thanks to further investigations, improved dating methods, and a diminished living memory of its history— it is thought to be less than 2,000 years old (see Marsh 2012). These two memory assemblages, formed at different times —with increased addends and distinct sets of archives— demonstrate how a single site can produce divergent archaeological narratives depending on the archives and addends assembled. A single archive is rarely sufficient for retrieval, and what is assembled with each archive is always of utmost importance.

We often restrict memory by postulating limits on the relevant addends and archives —this is typically the case when dealing with episodic human recollection of a specific occasion. Because episodic memory focuses on a particular collection of archives and addends, it suggests that when remembering is accurate, there is something like an identity between the

retained *x* and the retrieved *y*. It may seem that the encoding of the episode at the time of retention optimized retrieval, and that the retrieved *y* is factually identical to the retained *x*. This is what seems to be the case when I remember leaving my glasses on that table, and if the memory is correct, then the glasses were indeed left on that table. Consider now misremembering —or rather, the ways in which mistaken remembering is subject to correction. Misremembering, insofar as it uses a person's memory as an archive, is itself a memory assemblage —it brings together at least one archive and at least one addend that is taken to be innocuous. But when misremembering is diagnosed —either by the same person at a later time or by someone elseanother archive is introduced (and perhaps other addends as well). Suppose I recall eating corn tortillas with you in Granada, Spain, a decade ago, and you then tell me that you cannot remember whether they were corn or flour tortillas. You might add that corn tortillas were unlikely to be found in Granada at that time, and additionally point to an image of the restaurant menu we both remember. You thus correct me by offering another memory assemblage that appears more reliable because it draws on more archives. I might still acknowledge that, if we were merely concerned with our tortilla meal in Granada a decade ago, my initial recollection could be good enough. This brings us back to the significance of misremembering and its importance for understanding memory.

4. The cognitive import of misremembering revisited

De Brigard's model is one where hypothetical thinking geared towards action is what orients the endeavor of trusting something to memory. Although there are cases of misremembering that are genuine failures of the system —cases where what has been retrieved is duly corrected by engaging other retained archives— most common examples of misremembering can be explained by their relevance in future reasoning and counterfactual thinking (De Brigard 2014, p. 179). At least in people not diagnosed with pathological conditions, misremembering enables useful guidance to action. By focusing on action, reasoning, and thinking, De Brigard downplays the role of human episodic recollection as memory. Memory, he implies, should function with an aim that is not that of optimizing remembering. Cases of misremembering, thus, do not inform about how memory works but rather suggest that human memory systems are not proper memory systems —and are better not to be, given the unstable environment in which they evolved and the cost those systems would involve. Arguably, for De Brigard, there is no existing memory system that is geared towards memory —it would thus be a relief to find out that humans ended up endowed with the second best. As a consequence, misremembering could be as cognitively important as remembering, as it is also connected to a (psychological) system that enables counterfactual thinking. The cognitive import of memory is not that of somehow absorbing past and retained episodes and retrieving them, but rather that of thinking and reasoning hypothetically.

The diachrony associated with memory is accordingly of lesser importance since what matters is how future and counterfactual considerations enabled by remembering (and misremembering) have effects in action.

From the pan-mnemist point of view, misremembering is cognitively important because it may suffice to recall the correct elements that are often embedded within mistaken remembrances. That a case is one of misremembering becomes evident when we consider a broader collection of archives mobilized in a memory assemblage. Rather than assessing retrieval episodes solely in terms of their accuracy relative to what was retained, retrieval is also evaluated with reference to the relevant addends —the circumstances of retrieval, new information, other archives that have become pertinent, and so on. The mechanism behind misremembering is thus the mechanism of memory itself: the construction of assemblages that are sufficiently appropriate for the occasion. The process of drawing on available archives in an act of retrieval shaped by addends is identical to any other instance of employing memory, which is, according to the pan-mnemist view, an ubiquitous process. Memory assemblages are continually revised by new addends that enter the scene —addends that may bring additional archives to light and reconfigure what was previously retrieved.

To be sure, memory assemblages —through which retrieval takes place— are *situated*, in the sense that each act of retrieval is conditioned by the available addends. While De Brigard's model posits a cognitive ability that explains the frequency of misremembering, the memory assemblages model makes misremembering cognitively meaningful by describing it in terms of memory itself. Through the act of assembling memories, one seeks to make the best possible use of the available archives in order to produce a retrieval that responds adequately to the circumstances at hand. In terms of cognitive assessment, the resulting memory can be evaluated by the adequacy and sufficiency of the archives and addends mobilized. Yet such evaluations themselves rely on further assemblages —assemblages that foreground what is taken to be adequate and sufficient— and are likewise subject to correction. The entire process —including episodic future reasoning, counterfactual thinking, and the eventual corrections that may follow— is understood as a function of memory. Moreover, other cognitive abilities —not only thinking and reasoning— can be described in terms of memory assemblages. Take perception, for instance. In this model, what is stored in a perceived object —say, the dagger before me— is an archive that can be retrieved in an assemblage incorporating appropriate concepts, categories, or images, all of which are integrated with other relevant archives. Accordingly, misperception can be understood as a misassembled configuration of archives and addends, and its correction can likewise be described in terms of memory assemblages.

De Brigard's model, as I pointed out, bridges together memory and imagination —it does that under the umbrella of future reasoning and counterfactual thinking. When one is imagining, those latter capacities have to be at play, and this explains how memory and imagination are both potentially appropriate guides to action. The memory assemblages model, on the other hand, could picture imagining as assembling from imaginary archives (or addends) and then eventually mixing them with other memory assemblages. To be sure, those imaginary items need not be known to be so —one thing that imagination teaches us is that we are sometimes to discard what had been taken as an archive, at least for some purposes. Future reasoning and counterfactual thinking requires engagement of memory as one performs tasks related to do with hypothetical processing through archives and addends. On some occasions, what is imagined is good enough for such processing to come to appropriate conclusions —and guide action adequately. What the model enables is a continuity between memory and imagination through the process of (memory) assembling. Memory is not what it is —and it works in ways that are similar to imagination— because the world is too unstable or because it is too costly to retain all the details of a witnessed episode. Memory functions through assemblages because of the persistence of addends. It would thus make no sense to simply preserve content; retention is geared towards the addends that are coming —we remember in view of occasions that can appear with the upcoming addends. Because of the pan-mnemism associated with the model, it would also make little sense for a cognitive memory system to be other than a dynamic system of memory assemblages.

5. The retained and the retrieved content

I conclude with some remarks concerning the content of what is retained in memory —and its relation to what is ultimately retrieved in each act of remembering. Martin and Deutscher (1966) argued that memory traces are structural analogues of past events. Bernecker (2008) and others have proposed different accounts of memory traces. More recently, the very idea of memory traces has come under scrutiny (see Robins 2016 and 2020, for example). To be sure, traces in human memory are likely to be associated with dispositional properties in the brain that form units, like Semon's engrams (Semon 1909). Post-causalist theories (such as Michaelian 2016) replace traces with the activity involved in retrieval; they treat memory as a synchronic rather than a diachronic capacity. Retention is thus eliminated, along with the now-dispensed memory traces. In contrast, De Brigard (2020) defends the indispensability of postulating memory traces of some sort in order to explain certain memory phenomena. He argues that memory traces are instrumental in a causal account that elucidates the connection between the retained episode and the circumstances of recollection—there must be something that persists throughout the two different moments

that constitute the memory process. De Brigard brings to the fore an account of traces offered by Norman Malcolm (Malcolm 1977), according to which memory traces must (1) preserve the intentional content unchanged until recollection, (2) be structurally similar to what is remembered, and (3) play a causal role in the retrieval of the retained event. De Brigard argues that postulating a common underlying element present throughout the memory process —from retention to retrieval— is crucial for a causal account, even if Malcolm's first two requirements cannot be met by the memory traces he envisages. He agrees with Robins that the possible neural vehicles of memory traces are dynamically too changeable to retain content or display isomorphism with the state involved in retention, but he maintains that we cannot dispense with memory traces as an underlying element in the causal process of memory he takes to be adequate.

De Brigard (2020) claims that there ought to be an element that drives the action of memory between its two temporal poles —traces would ensure that memory is a process with duration, as they bridge the temporally discontinuous poles of retention and retrieval. If two people witness the same event but only one remembers it later, he finds it reasonable to infer that something was internally retained in one but not in the other. The explanation for this difference in recollection could lie in disease, a lack of attention, or neural architecture, but De Brigard insists that some kind of causal intermediary must be posited, and memory traces fulfill this role. In other words, something remains in one of the persons after the event witnessed by both, but not in the other. However, this stripping down of memory traces to whatever is permanent in the memory process —even if it retains no intentional content or displays no structural analogy with what is eventually retrieved— leaves us with a notion that is both too vague and geared only toward satisfying a perceived need for continuity between the two moments involved in memory. From the point of view of memory assemblages, memory traces can be understood as no more than archives, which may take any form —from neural configurations (or dynamics) to inscriptions on paper or wood. The content of the archive cannot be separated from the addends introduced in the process of retrieval. As a consequence, there is no structural analogy between the retained archive and the memory assemblage that is ultimately retrieved with the aid of further archives and addends. There is no underlying element linking the moment of retention to the corresponding memory assemblage, because retrieval in light of the addends will determine which archives are brought into the assemblage. Still, we can explain why two people who witness the same event may have different recollections: they retrieve through different assemblages, either because they have had different addends introduced since retention —for example, one of them has seen the event repeated many times afterwards— or because they access different archives —for instance, one of them had a marked experience during retention while the other retained something else in memory.

From the point of view of a broader account of memory, such as the one put forward by pan-mnemism, the issue is what precisely is the retained x —and, accordingly what is the retrieved y— in a memory assemblage. The archived content has itself the form of what is retrieved in the moment of retention —like encoding an experienced episode, writing a text, forming a landscape. In other words, if the retrieved content in a memory assemblage is a function from archives and addends, each archive is also a function of further archives and addends; while $y = f(x_1, \ldots, x_n, a_1, \ldots, a_n)$, each $x_i = f(x_{1'}, \ldots, x_{n'}, a_{1'}, \ldots, a_{n'})$.[1] From a pan-mnemist point of view, there is no room for an ultimate retained x; further, there is no content continuity and no structural analogue between x and y —after all y is a function of more than one archive and a collection of addends. It could happen that human (episodic) memory attempts to encode x, trying to prefigure what would be required in retrieval, but this is irrelevant for the process of assembling archives and addends. Nor, as De Brigard maintains (De Brigard 2014, p. 172, footnote), what is kept in retention has dispositional properties that are relevant for the memory process. Both x and y are archived for future use that can take various different shapes depending on the relevant addends at the time of retrieval. The content of each of these elements displays no general structure, no ontology apart from their nature as archives that could eventually be retrieved through methods that could even be unknown at the time of retrieval —archaeological archives are discovered after periods of time when they were unavailable or could not be inspected. In the process of assembling memory, there are no intermediaries apart from addends that are not anything constant, but which often make whatever was intended to be remembered less relevant.

From the pan-mnemist perspective, the content of what is retained can only be determined at the moment of retrieval —it is only within a memory assemblage that an archive acquires meaning or carries specific content. The retained content is just as dependent on the addends composing the assemblage as the retrieved content —both are shaped by the addends brought into play. If memory is understood as a ubiquitous phenomenon that relies on future assemblages of archives and addends, then the search for a general account of retained content is dissolved. In the case of human memory as a cognitive process, the issue shifts to one of brain archives or neural retention —that is, to the memory assemblages formed within neural networks in the brain. Here, the idea of assembling archives and addends becomes useful: neural networks combine not only with other networks but also with addends found in the relevant context. Although beyond the scope of this paper,

[1] It could be that the same archives (and addends) are used both in the assemblage $y = f(x_1, \ldots, x_n, a_1, \ldots, a_n)$ and in the assemblage $x_i = f(x_{1'}, \ldots, x_{n'}, a_{1'}, \ldots, a_{n'})$; that is, there could be a $x_i = x_{i'}$ (and some $a_i = a_{i'}$).

this suggests that a general architecture of memory as a pervasive phenomenon may provide useful tools for analyzing human memory processes. Furthermore, as I have argued here, this can be done without treating memory as a mere byproduct of some supposedly more encompassing cognitive activity.

References

Aristotle (1972). De Memoria et Reminiscentia. In R. Sorabji (Ed.), *Aristotle on Memory* (pp. 47–60). Brown University Press.

Bensusan, H. (2024). *Memory Assemblages: Spectral Realism and the Logic of Addition.* Bloomsbury.

Bernecker, Sven (2008). *The Metaphysics of Memory*. Dordrecht: Springer.

Brainerd, C. & Reyna, V. (2005). *The Science of False Memory*. Oxford University Press.

Clark, A. & Chalmers, D. (1998). The extended mind. *Analysis, 58*(1), 7–19. https://doi.org/10.1093/analys/58.1.7.

Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford University Press.

De Brigard, F. (2014). Is memory for remembering? Recollection as a form episodic hypothetical thinking. *Synthese, 191*(2), 155–185. https://doi.org/10.1007/s11229-013-0247-7.

De Brigard, F. (2020). The explanatory indispensability of memory traces. *The Harvard Review of Philosophy, 27*, 23–47. https://doi.org/10.5840/harvardreview202072328.

Heersmink, R. & Carter, A. J. (2020). The philosophy of memory technologies: Metaphysics, knowledge, and values. *Memory Studies, 13*, 416–433. https://doi.org/10.1177/1750698017703810.

Malcolm, N. (1977). *Memory and Mind*. Cornell University Press.

Marsh, E. (2012). The Founding of Tiwanaku. *Ñawpa Pacha, 32*, 69–187. https://doi.org/10.1179/naw.2012.32.2.69.

Martin, C. B., & Deutscher, M. (1966). Remembering. *The Philosophical Review*, *75*(2), 161–196. https://doi.org/10.2307/2183082.

Matthen, M. (2010). Is memory preservation? *Philosophical Studies, 148*(1), 3–14. https://doi.org/10.1007/s11098-010-9501-8.

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

Langland-Hassan, P. (2021). What sort of imagining might remembering be? *Journal of the American Philosophical Association*, *7*(2), 231–251. https://doi.org/10.1017/apa.2020.28.

Neter, J., & Waksberg, J. (1964). A study of response errors in expenditures data from household interviews. *American Statistical Association Journal*, *59*, 18–55. https://doi.org/10.1080/01621459.1964.10480699.

Noë, A. (2004). *Action in Perception*. The MIT Press.

Robins, S. K. (2016). Representing the past: Memory traces and the causal theory of memory. *Philosophical Studies*, *173*, 2993–3013. https://doi.org/10.1007/s11098-016-0647-x.

Robins, S. K. (2020). Stable engrams and neural dynamics. *Philosophy of Science, 87*(5), 1130–1139. https://doi.org/10.1086/710624.

Semon, R. (1909). *Mnemic psychology*. George Allen and Unwin.

Sutton, J. (2009). Adaptive misbeliefs and false memories. *Behavioral and Brain Sciences*, 32(6), 535–536. https://doi.org/10.1017/S0140525X09991488.

Tulving, E. (1983). *Elements of Episodic Memory*. Clarendon Press.

Vosgerau, G. (2010). Memory and content. *Consciousness and Cognition, 19*(3), 838–846. https://doi.org/10.1016/j.concog.2010.06.021.