(The paper was published in *Review of Philosophy and Psychology*, **11**, 487–505, <https://doi.org/10.1007/s13164-020-00461-1>. Please cite published version.)

Simulationism and the Function(s) of Episodic Memory

Arieh Schwartz

University of California, Davis

Abstract: According to simulationism, the function of episodic memory is not to remember the past, but to help construct representations of possible future episodes, by drawing together features from different experiential sources. This article suggests that the relationship between the traditional storehouse view, on which the function of memory is remembering, and the simulationist approach is more complicated than has been typically acknowledged. This is attributed, in part, to incorrect interpretations of what remembering on the storehouse view requires. Further, by appeal to function pluralism, the article questions both the assumption that the traditional view and simulationism are inconsistent, and the simulationist’s inference to the best explanation strategy that is based upon this assumption. The article then provides an evaluation of the simulationist argument against the traditional view, and finds that it in need of further support.

This is a moment of excitement and upheaval in the study of memory. The old picture of memory as a faculty for remembering—storing and retrieving facts and experiences—is being superseded in the eyes of many researchers. Taking its place, is a new picture on which the function of episodic memory[[1]](#footnote-1) is not to remember past experiences, but to help construct representations of possible ones, by assembling features drawn from many experiential sources (see, e.g., Suddendorf and Corballis, 1997, 2007; Schacter and Addis, 2007; Boyer, 2009; De Brigard, 2014a; Michaelian, 2016). “Smulationist” views, as I’ll call them, are a motley crew. They are, first and foremost, hypotheses about the function of episodic memory, in one sense or another of the term “function.” The “function” of episodic memory, variably in an etiological, causal-role, or survival-value sense, is supposedly not to remember the past, but to imagine or simulate hypothetical episodes of experience.

The excitement surrounding simulationism is due, at least in part, to the revisionism of its claims. This excitement has been beneficial for the study of memory. Nevertheless, while simulationism is a recent paradigm that should be nurtured, its revisionary claims should not be accepted unreflectively. The goal of this article is not to establish whether simulationism or a traditional view about the function of episodic memory is correct. The goal is rather to investigate the relationship between these approaches. I submit that it is more complicated than previously acknowledged. It is important to separate two threads of the simulationist story. There is the negative project, which argues that episodic memory *is not* for remembering. Then there is the positive project, which argues that episodic memory *is* for simulating. This article outlines a challenge to the negative project, and its use in motivating the positive project. The challenge has broader relevance, however, in promoting due caution in drawing epistemic conclusions from false-memory research. The article will proceed as follows. Part one explores the traditional idea that memory is for remembering. Part two aims to accurately exposit the historical developments and arguments that motivate simulationist accounts. Part three offers critical discussion. Part four concludes.

**Part 1. The Storehouse of Memory**

 The aim of the negative project is to discredit the old idea that the function of memory is remembering. “[W]e may,” according to a prominent simulationist, “have just been wrong about the mind…remembering does not seem to be what memory is for,” (De Brigard, 2014a, p. 168). Since simulationism concerns the specialized topic of episodic memory, one might wonder whether we, the folk, hold views on it to be corrected. But surely, while the episodic/semantic distinction may be recent (Tulving, 1972) and specialized, memory *in general* has been widely assumed to serve the remembering function. So, if a genuine memory capacity is not for remembering, then folk-psychology’s general assumption about memory function is flawed.

Simulationists can be forgiven for not providing a detailed analysis of the traditional notions of memory and remembering; no easy task. But to argue that memory is *not* for remembering will inevitably involve some assumptions about what remembering requires. Indeed, the negative project appears to rely on the assumption that remembering requires absolute or perfect sameness between the representations deposited in and retrieved from memory. This assumption is, I believe, less than charitable to the folk view, for reasons that will soon become clear.

While the traditional notions of memory and remembering have obviously not been codified, perhaps the common wisdom can be discerned in canonical memory metaphors. Foremost among them are the waxen block, and the storehouse. Endorsing the waxen block metaphor that Plato had introduced in *Theatetus*, Aristotle refers to memory as “a sort of picture…for the movement produced [by perception in the soul] stamps almost a sort of impression of the sense-impression, similar to what is done by people using their seals,” (450a25-32). A related orthodox idea holds that memory is a sort of *repository* for images. Augustine, for example, famously compared memory to a “storehouse,” a “vast palace,” a “vast hall,” and a,

huge cavern [that] receives all these perceptions, to be recalled when needed and reconsidered... The objects themselves do not enter, but the images of the perceived objects are available to the thought recalling them. (p. 484)

The main idea of these metaphors seems to be that experience generates representations with both intentional and physical aspects, such that access to the former can be preserved by means of the preserving the latter.[[2]](#footnote-2) When the process *succeeds*, when remembering takes place, the representation retrieved accurately reflects the relevant past experience, because it was, in a meaningful sense, causally produced or even “imprinted” by it. On the other hand, *failures* to remember can be explained by appeal to the causal dynamics of imprinting, storage, and retrieval.[[3]](#footnote-3)

 The general storehouse metaphor needs to be unpacked to be evaluable. Of particular interest, is how *remembering* should be understood within this framework. We can plausibly consider “remembering” as a success term to be applied when, (given suitable background conditions,) what is retrieved from the storehouse is relevantly “the same” as what was deposited. But remembering, the success of the storehouse, might be measured by different standards of “sameness.” It is very important to get clear on which standard is actually employed in remembering judgments.

One proposal is that the storehouse of memory succeeds only if the *individual* physical representation that is retrieved from the storehouse is numerically identical to the one that was deposited, (and continuously stored).[[4]](#footnote-4) A pair of observations weigh against this proposal. First, it seems perfectly coherent, given the storehouse metaphor, to think of what is stored as an *ability* to represent the relevant episode, rather than as a particular physical representation itself.[[5]](#footnote-5) On this view, remembering would occur when (among other things) the relevant abilities are preserved, although the particular physical representations that were input to memory are not. So, remembering, in the storehouse framework, *need* not involve identity between the representational vehicles deposited/retrieved. Second, remembering, in the storehouse framework, *should* not be understood to depend on such an identity relation. Remembering judgements seem quite unbeholden to facts about which individual physical representations are involved in recollection. We don’t usurp the neuroscientists in making claims of the form, “*S* remembers episode *E*.” Numerical identity between the particular internal representations that are memory’s input and output should not be imposed as a requirement of remembering.

Another proposal views sameness not of content *bearers*, but of content itself as criterial for remembering. This proposal comes in different versions. On one version, a subject only counts as remembering an episode if the representations deposited and retrieved have “the same” content, in the sense of sharing *all* content properties. This standard also appears too strict,[[6]](#footnote-6) in light of the observation that remembering is not ruled out by differences, even big differences, between the contents of the representations that are deposited in and retrieved from memory. Suppose, for example, that Jones attempts to episodically remember leaving home this morning in order to verify that he locked the front door. Jones may remember the episode, even if the detail that it had been raining was not preserved from the representation that was deposited to the representation that is retrieved. Let’s call the idea that correct remembering attributions tolerate differences between the contents of representations deposited in and retrieved from memory, *The Tolerance Principle*,(or *Tolerance*, for short).

In light of *Tolerance*, a modified content standard may assert that a subject remembers an episode only if the representations deposited in and retrieved from memory are the same insofar as they share *most* of their content features; (that way it will not matter if they diverge with respect to a few). Personally, I am uncertain how one would go about tallying up *all* the features of a representation in the way this proposal requires. But even if an adequate counting scheme were discovered, the proposal would still neglect an important fact. In determining whether a subject remembers an episode, it will matter very much whether the episodic representations deposited and retrieved are the same with respect to some features of content, it will matter very little whether they are the same with respect to other features of content, and the subject’s epistemic context will be what makes the difference.

Remembering judgments are *tolerant* because they are context-sensitive. Whether or not a subject counts as remembering an episode crucially depends on the epistemic use(s) to which the information retrieved is to be put. For example, suppose Jones merely wishes to verify that the front door was locked this morning. He may be correctly judged to remember the episode, even if the content of the representation retrieved is *as of* locking the deadbolt, while the content of the representation deposited was *as of* locking the door-knob. However, the same instance of recollection might count as *misremembering*, if Jones’ purpose were to determine whether he had *only* locked the deadbolt. The familiar point I hope this makes clear is that the purposes of episodic recall highlight the dimensions of the target episode that need to be represented, and the degree of precision they need to be represented *with* in order for the subject to count as remembering the episode in question.[[7]](#footnote-7)

In general, (and assuming background conditions are met,) episodic remembering may occur when the representations deposited in and retrieved from the storehouse have the same content in the relevant dimensions, at the relevant levels of description. There are various ways this idea could be articulated. One suggestion is that remembering requires that the representations deposited and retrieved are “the same,” in the sense of being *tokens* of the same contextually-relevant state *type*. Remembering, on this view, would require that the concrete particular representations deposited in and retrieved from memory instantiate relevantly the same type of representation given the epistemic context, insofar as they instantiate the same relevant content property. This proposal allows for the *tolerance* of remembering judgments in the same way that ‘tiger’ and ‘TIGER’ count as tokens of the same word type, despite their various differences. Moreover, context may affect whether remembering occurs in much the same way that judged by word *type*, ‘tiger’ and ‘table’ count as tokens of different types, but judged by part of speech *type*, they are tokens of the same type. There are various other ways the storehouse metaphor could be developed.[[8]](#footnote-8) The important thing to emphasize is simply that, however it is developed, it should accommodate *tolerance*. For remembering, content need only be preserved to the extent that representations retrieved are “the same” as representations deposited in the respects and at the levels of description relevant to the tasks at hand.

 While the storehouse view is rooted in old memory metaphors, it is not a mere archaic curiosity. The “causal theory” of memory, (Martin and Deutscher, 1966) which still has adherents in analytic philosophy (see, e.g., Bernecker, 2010; Cheng and Werning, 2016) is arguably a version of a storehouse view.[[9]](#footnote-9) According to the causal theory, a subject remembers something when (roughly) her experience of it caused an internal state or contiguous lineage of states that causally explain why the representation of it she retrieves is accurate (to the extent that it is).[[10]](#footnote-10) The causal theory was originally intended as a conceptual analysis or descriptive definition of what it means for a subject to remember, or to have a memory (Martin and Deutscher, p. 190). Moreover, as Koriat and Goldsmith (1996) point out, the storehouse metaphor also seems to underlie the quantificational tradition of testing that has been popular in memory science since Ebbinghaus.[[11]](#footnote-11) Their point is that research that proceeds by counting the number of items from study that are recalled at test assumes that memory is a sort of storehouse into which representations can be deposited, and from which, retrieved. So, it is fair to say that diverse strands of philosophy and memory science remain aligned with the basic storehouse intuition.

**Part 2: The Rise of Simulationism**

One of the major discoveries of memory science has been that memory is not a unified phenomenon. Rather, there are different functional kinds of memory, that are at least partially dissociable.[[12]](#footnote-12) Simulationism is a hypothesis about episodic memory. The term episodic memory was introduced in contrast to “semantic,” or merely factual memory (Tulving, 1972, p. 384).[[13]](#footnote-13) The original distinction had several components. The primary component of the distinction was that episodic memory includes, (while semantic memory lacks,) details of the temporal context in which the relevant information was acquired.[[14]](#footnote-14) The distinction gained support from later work with the amnesic patient, K.C., which suggested that the system responsible for (broadly) memories of personal experiences could be incapacitated, while the system responsible for memories of factual information was not.[[15]](#footnote-15) Disagreement persists on how to draw the distinction between these cognitive systems and their instances precisely. On one approach, episodic memories are typified by including contextual details about where and when the represented event occurred, (Cf. Tulving, 1983, p. 25). A criticism of this approach is that it cannot adequately distinguish episodic from semantic memory, because semantic memories may include relevant contextual details, while episodic memories may lack them. On a different approach, while episodic memory *may* include contextual details, it is distinguished by a peculiar “autonoetic” or “self-knowing” phenomenology as of having personally experienced the represented event in one’s past (Cf. Tulving, 1985, 1993, 2001).[[16]](#footnote-16) A problem faced by the view is that it hard to operationalize; it is hard to know when a creature is having one phenomenology rather than another, or rather than none at all. Simulationists are not univocal in how they characterize episodic memory.[[17]](#footnote-17) While unclarity about the nature of episodic memory may raise difficulties confirming the hypothesis that its function is to help simulate hypothetical episodes, because my focus is on the negative project, I leave this issue aside.

 The negative and positive projects are not always distinguished. This is because the simulationist view is usually argued for on the basis that, compared against the traditional remembering view, it provides the best explanation of two bodies of empirical evidence.[[18]](#footnote-18) (The fairness of this strategy is contested in part three.) The first body of evidence consists in false memory research that suggests that episodic representations are often inaccurate, presumably because they are constructed at the time of retrieval. The second body of evidence consists in empirical correlations between episodic memory and episodic simulation of possible future events. While the first body of evidence is mainly used to advance the negative project, the second body of evidence is mainly used to advance the positive project. Nevertheless, since simulationism and the storehouse view are presented as alternative hypotheses in an inference to the best explanation, the success of the negative project in demonstrating that the storehouse view is incorrect is taken to support the positive view, that episodic memory is for simulating possible episodes.

 A lesson often drawn from false memory research is that the storehouse view is incorrect. Sir Frederic Bartlett (1932) was an early exponent of this opinion. However, Bartlett’s interpretation of the storehouse view was overly strict. Objecting to Head (1920)’s description of “the cortex as ‘a storehouse of past impressions’,” Bartlett remarks that “a storehouse is a place where things are put in the hope that they may be found again exactly as they were when first stored away,” (p. 200). Bartlett distinguished between *reproduction* and *reconstruction*, and assumed that the storehouse view requires the former. “Reproduction” refers to “literal recall,” achieved through “re-excitement” of an “individual trace,” that had been imprinted by past experience and safely stored away (p. 204). Far from “literal recall,” Bartlett found that subject’s memories featured omissions, as well as distortions shaped by current interests and point-of-view. “In the many thousands of cases of remembering which I collected, a considerable number of which I have recorded here, literal recall was very rare… [The] re-excitement of individual traces did not look to be in the least what was happening,” (p. 204). Bartlett reasoned that such errors are hard to explain given the (supposed) commitment of the storehouse model: that recall is produced by the re-excitement of the individual trace that was imprinted by experience.[[19]](#footnote-19) On the picture Bartlett developed, remembering is rather *reconstructive*, beginning with a general “attitude” or feeling in response to stored traces, which serves as the basis around which a representation is presently built up (p. 207-214).[[20]](#footnote-20)

While the false memory tradition Bartlett began has usefully dispelled the notion of reproductive memory, the storehouse picture of remembering does not require reproduction. In reproduction, the same “individual trace” is re-activated, and “literal recall” results. But, (as the previous section argued,) the storehouse view requires neither that the individual representations deposited and retrieved are numerically identical, nor that their contents are. Considered as a storehouse, memory is closer to a bank than a daycare; it is of little moment whether the items deposited and retrieved are the exact same ones.

The tradition of false memory research has remained remarkably active. New paradigms have revealed surprising ways in which memory is inaccurate, and/or manipulable. They also suggest, although they do not prove, that our *own* memories may be inaccurate in these interesting ways. The *DRM*, and *misinformation* effects are among the most dramatic findings. I will explain them briefly, before discussing their significance to the negative project. The *DRM* effect gets its name from a paradigm developed by John Deese (1959), and resurrected by Henry Roediger and Kathleen McDermott (1995). In this paradigm, subjects are given study-lists of items, (words, images, etc.,) that are semantically related to a further unpresented item, (the “critical lure”).[[21]](#footnote-21) At test, subjects tend to falsely recognize or falsely recall critical lures at rates comparable to those exhibited for items actually encountered in the study phase. It is widely seen as significant that even list-learning, presumably a highly reliable form of memory, is prone to false recall and recognition errors.[[22]](#footnote-22)

Research in the *misinformation paradigm* pioneered by Elizabeth Loftus and colleagues found that subjects can be induced to recollect events in distorted ways depending on what information they are presented after the fact. For example, in a classic study, Loftus and Palmer (1974) report that subjects were more than twice as likely to remember seeing broken glass in a car-crash video if the accident had been previously described with the verb “smashed” rather than “hit.” Moreover, other studies have found that false recall can be induced for events that never even occurred (see, e.g., Loftus and Pickrell, 1995).

This and other false memory research[[23]](#footnote-23) suggest that our memories *may* often be inaccurate in fascinating ways. Of course, it is a controversial leap from the evidence that memory errors can be observed/induced in laboratory settings, to the conclusion that our memories exhibit these sorts of inaccuracies in everyday life. Nevertheless, such research may justifiably raise our suspicions that memory is more prone to inaccuracy than previously acknowledged. What the evidence does not indicate, and this is crucial to the negative project, is that remembering is relatively infrequent compared to misremembering. The mere fact that a subject retrieves an episodic representation that is inaccurate in some respects at some level of description does not imply that she is not remembering the episode. What matters is whether the inaccuracies of the representation prevent it from doing the real-life epistemic work it is needed for. It is unclear how frequency data could be gathered on this. But this is the sort of data that simulationists would need to claim that remembering is a comparatively scarce occurrence. (We will revisit this theme in part three.)

The motivation of simulationism is to explain why we have an episodic memory system that is constructive, that seems to assemble content drawn from different sources, and that is thereby prone to inaccuracy in the ways that false memory research suggests. For example, cognitive neuroscientists Daniel Schacter and Donna Rose Addis (2007) first introduce their view, the *constructive episodic simulation hypothesis*, in the following way.

We consider some recent work concerning the neural basis of memory construction with a view to addressing a question concerning its function: why does memory involve a constructive process of piecing together bits and pieces of information, rather than something more akin to a replay of the past? Several researchers have grappled with this issue and proposed various reasons why human memory, in contrast to video recorders or computers, does not store and retrieve exact replicas of experience. (773-774, references omitted.)

This passage illustrates the aim to provide an adaptive explanation for why episodic memory is not reproductive, “does not store and retrieve exact replicas of experience,” for why episodic memory is error-prone. The relevant assumption seems to be that if accurate remembering were episodic memory’s (sole or primary) function, then episodic memory would be less prone to inaccuracy. Therefore, accurate remembering is not episodic memory’s (sole or primary) function. This reasoning is what I have called the negative project.[[24]](#footnote-24) Thomas Suddendorf and Michael Corballis (2007), for example, pursue it as follows.

Although episodic memory preserves something of the particularities of individual events, it is often unreliable and subject to distortion, as we have seen…The fact the episodic memory is fragmentary and fragile suggests that its adaptiveness may derive less from its role as an accurate record of personal history than from providing a ‘vocabulary’ from which to construct planned future events (and perhaps to embellish events of the past),” (302-303, references omitted).

The negative project is developed most thoroughly in Felipe De Brigard’s (2014) article “Is Memory for Remembering?” His response, in short, is No. The target of the inference is “the natural conclusion that the function of memory is to store, retain, and then to reproduce (or make available, or reconstruct) the contents of past experiences at a later time,” (p. 158, references omitted). The inference begins by pointing out that given this storehouse sort of view,

[when] memory fails either to faithfully store or to bring back to mind a past experience with fidelity…memory fails to perform its function. Despite its intuitive appeal, saying that false and distorted memories are a failure of memory may force us to accept that we have a memory system that regularly and systematically malfunctions. Evidence gathered by cognitive scientists in the last four decades makes it clear that false and distorted memories are a common occurrence in our daily lives. (p. 159)

At this point one may naturally wonder how can the received view defend the claim that false memories are the product of a malfunctioning faculty in the false of their pervasiveness and regularity. Furthermore: why would we have a system that malfunctions so constantly and so systematically? (p. 163)

De Brigard’s conclusion is that episodic memory’s function is not to remember, but rather to contribute to a larger cognitive system responsible for simulating hypothetical episodes of experience.

 The positive project is advertised as the best explanation of empirical correlations between episodic memory and episodic simulation of possible, e.g. future, experiences. The correlations come in a few varieties. There are: 1) neuroimaging correlations; 2) “paired deficits”; 3) developmental correlations; and 4) phenomenological correlations. The most important correlation for the simulationist’s case was reported in neuroimaging studies from the early 2000’s. Namely, areas of brain activation observed during episodic memory tasks are also observed during tasks involving simulation of future events (see, e.g., Atance and O’Neil, 2001; Okuda et al., 2003; Hassabis and McGuire, 2007; Szpunar et al., 2007; Schacter et al., 2007). This discovery was also the historical catalyst for the simulationist program. Importantly, patterns of activation in the two types of tasks were not fully similar. Episodic future simulation tended to involve activation in a larger network of brain areas than episodic recall.[[25]](#footnote-25) Nevertheless, a “core brain network”[[26]](#footnote-26) seemed to be involved in both types of tasks. From this, simulationists infer that episodic memory is probably a part of this larger cognitive system. The operative inference is not fully clear. It appears to conclude that the psychological process of episodic memory is probably involved in the broader psychological process episodic simulation, because the areas of brain activation observed during the psychological process of episodic memory overlap a larger network of brain activation observed during episodic simulation tasks. The cogency of this inference is a topic that will have to wait until another occasion.

 The second correlation simulationists cite is the paired-deficit neurological patients exhibit between remembering episodes from the past and imagining episodes that might occur in the future. For example, Tulving (1985) reported that N.N. (also known as K.C.) was not only unable to remember past experiences, but also to imagine future ones; famously describing his mental state when attempting either as involving, “the same kind of blankness,” (p. 4). A similar paired-deficit was reported Klein et al. (2002) working with patient G.B., and by Hassabis et. al (2007) working with a group of amnesic patients.

 Third, simulationists sometimes point to a developmental correlation between episodic memory and episodic simulation. Both capacities appear to emerge around the same time in development, between 3-4 years old (Busby and Suddendorf, 2005; Suddendorf and Corballis, 2007).

 Fourth, simulationists often point to phenomenological correlations between episodic memory and episodic future simulation. An often cited example is the finding of D’Argembeau and Van der Linden (2004), that subjects report a decrease of phenomenal richness in their representations of both past and future events that is proportional to the (supposed) remoteness of these episodes from the present.

 Simulationists take these data to suggest that episodic memory is a component of a larger cognitive system responsible for simulating possible experiences. For example, Kourken Michaelian (2016) reports that,

episodic memory is currently viewed, by most psychologists working in the area, as one instance of a more general capacity allowing the agent both to re-experience past episodes and to “pre-experience” possible episodes … The standard view, in short, is increasingly that episodic memory is one function of a more general episodic construction system, a process not different in kind from imagining a range of nonactual episodes. (p. 98-99)

The simulationist hypothesis is that the reason we have constructive episodic memories, why episodic memory is error-prone, is so it can play its part in the evolutionarily more important task of simulating future episodes. Suddendorf and Corballis, for example, observe that “there is a growing recognition that mental time travel into the past [i.e. episodic memory] and future are related, and that the ultimate evolutionary advantage must lie with the capacity to access the future,” (2007, p. 299). Simulationists are not univocal about the precise role of episodic memory in simulating possible events. Some hold that “the primary role of mental time travel into the past [i.e. episodic memory] is to provide raw material from which to construct and imagine possible futures,” (Ibid., p. 302). Others hold that, in addition to providing material, episodic memory is itself part of the instrument that constructs simulations.[[27]](#footnote-27)

Recent work attempts to directly address the question of how simulating future events yields evolutionary advantages. Pascal Boyer (2009), for example, argues that in making future rewards/punishments feel current, simulating the future lessens the extent of “temporal discounting,” and thereby improves our planning/decision-making. Relatedly, Christoph Hoerl and Teresa McCormack (2016) suggest that simulation improves decision-making by allowing agents to pre-experience the regret that poor decisions might bring about. While the thought that episodic memory contributes, in some way, toward episodic simulation is plausible, the conclusion that episodic memory does not have the function of remembering the past has not been adequately argued.

**Part 3. Evaluating the Negative Project**

 While my aim is merely to evaluate the negative project, one general complaint is in order. Simulationist accounts aim to determine the “function” of episodic memory without being precise about what is meant by the term. This makes it difficult to interpret both the negative and positive projects. In what sense of “function” is the function of episodic memory not to remember past experiences? In what sense of “function” is the function of episodic memory to simulate possible experiences? It is unclear.

 There is a large literature in philosophy about how to understand what “function” means in scientific theorizing. Moreover, since Millikan (1989), there has been growing support for the idea that there is no single correct conception of ‘function,’ which rather takes on different but complimentary senses in different forms of scientific explanation.[[28]](#footnote-28) At least three main conceptions of function should be distinguished. First, there are *causal-role functions*, ascribed to explain how a device works now. The approach was pioneered by Robert Cummins (1975). On this view, a function is ascribed to a device within the context of a functional analysis of how the system containing the device, through the “programmed” execution of simpler component steps, produces a complex product of interest. The function of the device is its causal contribution to this step-wise production.[[29]](#footnote-29) Second, there are *etiological* or *selected-effects* functions, ascribed to explain the “purpose” of a device in the sense of *why it is here.* This approach was pioneered by Wright (1967), Millikan (1984), and Neander (1991). On this view, the function of a device is roughly the sort of performance that forerunners of the device gave in the past which (partially) explains why they were continually copied, and why, in turn, the current device is here.[[30]](#footnote-30) Third, there are *survival-value* functions (Tinbergen, 1963; Godfrey-Smith, 1994), ascribed to explain the contribution of a device to the survival (as a proxy for the fitness of) the organism. Ascribing a survival-value function to a device could address different questions about it. For example, employing the *modern history* conception of function (Godfrey-Smith, 1994), one could aim to identify a selection pressure that has been operative on the character. On this view, the function of a character is a disposition in virtue of which it was maintained during the most recent history of selection. Whereas, employing the *propensity function* view (Bigelow and Pargetter, 1987), on which the function of a character is a propensity for selection it confers on the organism, one could aim to predict the form the character is likely to have in future generations.[[31]](#footnote-31)

Two claims made earlier can now be justified. First, the sense of “function” assumed in simulationist accounts is not always clear. For example, Schacter and Addis initially describe their account of episodic memory as,

addressing a question concerning its function: why does memory involve a constructive process of piecing together bits and pieces of information, rather than something more akin to a replay of the past?... [Why does] human memory, in contrast to video recorders or computers…not store and retrieve exact replicas of experience…We focus on one hypothesis concerning the origins of a constructive episodic memory. (p. 773-774, references omitted)

The etiological approach to function is the most relevant to answering questions about the origins of biological systems. Somewhat later, however, Schacter describes the simulationist project differently.

A central tenet of the constructive episodic simulation hypothesis (Schacter & Addis, 2007) and related perspectives (Suddendorf & Corballis, 2007) is that the ability to flexibly recombine elements of past experience into simulations of novel future events is an adaptive process, sufficiently beneficial to the organism that it is worth the concomitant cost in memory errors that result from occasionally mistakenly combining those elements. (2012, p. 606)

[N]ote that I use the term “adaptive” in this article to refer to a beneficial characteristic of an organism, and make no claim about the evolutionary origins of adaptive constructive processes. (2012, p. 604)

In this context, the *constructive episodic simulation hypothesis* appears to ascribe a sort of survival-value function, and not an etiological one.[[32]](#footnote-32) Unclarity about the precise sense of function at issue makes it difficult to evaluate the simulationist project, as different types of functional ascription must rely on different types of evidence.

 Second, I have claimed that the negative and positive projects should be kept distinct. Simulationists do not usually separate them, because they argue for simulationism based on an inference to the best explanation. It is peremptory to set things up in this way. For cogent inferences to the best explanation usually require that the hypotheses considered are both exclusive and exhaustive. But this has not been shown. We cannot simply assume that the traditional remembering hypothesis and the simulationist hypothesis are inconsistent alternatives. For one thing, it is not clear whether simulationism and the storehouse view even invoke the same conception of function. While simulationism often seems to be an etiological view, the traditional storehouse view was not historically put forward with evolutionary considerations in mind.[[33]](#footnote-33) If these hypotheses assume different conceptions of function, we can’t assume they are inconsistent. For a system may perform different “functions,” given different conceptions of function.[[34]](#footnote-34) Moreover, even holding the conception of function fixed, etiological, causal-role, and survival-value conceptions of function each allow for the possibility that a device has multiple functions. (Different effects may explain why a device was continually copied, and eventually why it is here. A device may make different causal/mechanistic contributions to the production of different complex products. And a device may have promoted, or may currently promote survival in different ways.) It is really an empirical matter whether a device performs multiple functions or not. It should also be remembered that for the inference to the best explanation for simulationism to succeed, all plausible alternative hypotheses regarding episodic memory function would have to be considered.

With this background in place, it is time to assess the negative project itself. The argument can be reconstructed as follows. Each premise will be considered in turn.

*Negative*:

P1) False memory research shows that episodic memory is often inaccurate in various ways.

P2) If the storehouse view is correct, i.e. if the function of episodic memory is to allow the subject to remember the past (accurately), then (from P1) episodic memory malfunctions pervasively.

P3) It is implausible that episodic memory malfunctions pervasively.

C) The function of episodic memory is not to allow the subject to (accurately) remember past events. The traditional view is incorrect.

According to premise one, false memory research suggests that, in the general population, episodic memory is often inaccurate. False memory research cited by simulationists does not come close to *proving* this. The epistemic significance of empirical findings for determining the nature of mental states, e.g. memory, depends on metaphilosophical assumptions. There are ways of individuating mental faculties on which false memory research does not reveal much about episodic memory *per se*. A traditional philosopher might take it that memories are *necessarily* accurate (see, e.g., Cheng and Werning, 2016). It might, for example, be claimed on *a priori* grounds that a subject can only have a memory of what was the case. For this sort of thinker, false memory research would only suggest that subjects are prone to mistake various other mental states for memories, and neuroimaging dta would only suggest that genuine episodic memories and ersatz episodic memories are subserved by the same brain areas. Even allowing that genuine memories can be inaccurate, however, it remains unclear how pervasively inaccurate episodic memory really is, (for reasons discussed in part two). It is unclear how to draw inferences from findings observed/induced in the laboratory setting to conclusions about how frequently our episodic memories are inaccurate in everyday-life. It seems that the most we can say with high confidence is that the false memory research suggests that our episodic memories are *sometimes* inaccurate, to *some* degree, and in *some* respects. For argument’s sake, however, let’s grant the premise, as written.

Premise two is the claim that given false memory research, were the storehouse view correct, memory would malfunction pervasively. Note, however, that this premise only follows if a strict conception of the storehouse view were obligatory. For example, if on the storehouse view, remembering required “reproduction,” in the sense of Bartlett, then any deviation from “literal recall” would amount to misremembering—a malfunction of memory. But the storehouse view should not establish perfect recall as a condition on remembering, (as argued in part one). What matters for remembering is that the content of the representation retrieved from the storehouse is the same as the content of the representation deposited there in the dimensions and at the level of description that are relevant to the demands of the retrieval context. It makes little sense to classify a subject as remembering or misremembering an episode unless we know about the retrieval context, and the reasons and objectives of the subject. Pervasive episodic memory inaccuracy only suggests pervasive misremembering, (and therefore pervasive malfunction, on the traditional view,) if these pervasive inaccuracies prevent episodic memories from successfully fulfilling their contextually-specific roles. So, while false memory research may suggest that episodic memory is inaccurate in interesting ways, it tells us little about the frequency of remembering and misremembering. For that, we would need to know whether episodic memory inaccuracies mattered in the contexts of retrieval. And, it is far from clear how this data could be gathered. For the sake of argument, however, let’s assume that false memory research shows that misremembering is pervasive, and that episodic memory pervasively fails to achieve its function of perfect recall.

 Premise three is the claim that it is implausible that episodic memory malfunctions pervasively. Presumably, this claim is not supposed to rest on the idea that episodic memory is some uniquely well-functioning system. Rather, it seems that premise three is supposed to rest on some more general principle about the nature of functions, in one sense or another of “function.” Given that it is unclear what type of functional explanation simulationists intend to provide, there are (at least) three versions of the missing general principle that could be assigned the relevant philosophical work.

*R1*: For any device, it is implausible that it pervasively fails to achieve its *causal-role* function(s).

*R2*: For any device, it is implausible that it pervasively fails to achieve its *etiological* function(s).

*R3*: For any device, it is implausible that it pervasively fails to achieve its *survival-value* function(s).

These principles will be considered in turn. None of them appear very plausible.

 Principle R1 seems arbitrary. It is unobvious why it would be implausible for a device to fail to achieve one of its causal-role functions pervasively. Recall that ascriptions of causal-role functions are relative to functional analyses or explanations. This feature is built into Cummins’ classic definition.

*X* functions as a *φ* in *s* (or: the function of *x* in *s* is to *φ*) relative to an analytical account *A* of *s*’s capacity to *ψ* just in case *x* is capable of *φ*-ing in *s* and *A* appropriately and adequately accounts for *s*’s capacity to *ψ* by, in part, appealing to the capacity of *x* to *φ* in *s*. (1975, p. 726)

While some analyses may be more appropriate than others, there is no rule that we can only ascribe causal-role functions to systems that would not fail to perform these functions pervasively. Remarking on the analysis-relativity of his account, Cummins writes,

When a capacity of a containing system is appropriately explained by analyzing it into a number of other capacities whose programmed exercise yields a manifestation of the analyzed capacity, the analyzing capacities emerge as functions. Since the appropriateness of this sort of explanatory strategy is a matter of degree, so is the appropriateness of function-ascribing statements. (Ibid., p. 765)

There is no obvious reason it would be categorically *inappropriate* to explain a product that is rarely produced by way of functionally decomposing the system causally responsible for it. Thus, *R1* appears unmotivated.

 Another potential way to support premise three is to invoke the principle *R2*, that it would be implausible for any device, (biological device, anyway) to fail to achieve its etiological function pervasively. Principle *R2* may seem likely in light of how apparently well-designed many traits are. Perhaps *R2* could be motivated along the following lines. Natural selection, on the whole, presumably “chooses” organisms with the best performing variants of traits. Whereas, organisms with worse performing variants were presumably less fit: hence, survived less long and left fewer offspring. Therefore, it is implausible that organisms would have traits that pervasively malfunction in the etiological sense. While this way of thinking may contain a kernel of truth, we should not be too hasty to accept principle *R2.* Recall that if we grant that remembering requires perfect accuracy, (and it doesn’t,) then we are forced to say that episodic memory pervasively malfunctions the only in sense of pervasively failing to perform its function *perfectly*. Accuracy, after all, is a matter of respect and degree. The false memory research by no means suggests that wildly inaccurate confabulations are the usual product of episodic memory. But there are numerous reasons traits may not perform their etiological functions *perfectly*. One reason is that natural selection does not “build” traits from scratch, but rather makes use of the materials that are available. Gould (1980) famously made this point in reference to the giant panda’s “peculiar thumb,” which is an enlarged wrist bone (*radial sesamoid*). Given the limitations of the building material, the thumb is a rather “clumsy contraption, not a lovely contrivance,” with respect to its etiological function of helping the panda to grasp and strip bamboo (p. 24). A second reason traits may exhibit sub-optimal performance with respect to their etiological functions is that they often do not needto perform their functions well. Being good enough, and heritable may suffice for a trait to persist in a lineage. Assuming there must be an ulterior adaptive reason why we have only an imperfect episodic memory instead of something more akin to a video-camera is like assuming there must be an ulterior adaptive reason why we have spines that don’t evenly distribute our weight.[[35]](#footnote-35)

 Lastly, premise three might draw support from *R3*, the principle that it would be implausible for a device to fail to achieve its survival-value function pervasively. Perhaps the thought would be that organisms possessing devices that pervasively fail to achieve their survival-value functions generally fail to survive. Principle *R3* is difficult to evaluate unless what is meant by a survival-value function is made precise. Assuming Godfrey-Smith’s modern history account, the principle becomes: it is implausible for a device to pervasively fail to perform the function for which it was maintained during the most recent history of selection on the trait. But this principle is not plausible. For example, natural selection may not be presently occurring, so failure of a trait to perform its survival-value function may not matter. Alternately, the trait’s modern-history function may no longer be important to survival or reproduction given the current environment. Moreover, even assuming the trait is currently under selection, an organism with a pervasively malfunctioning trait may survive and reproduce if the trait performs well enough. For example, one successful performance of a predator avoidance device may compensate innumerable false negatives. Or the organism may survive if the poor performing trait can be compensated for by other traits, or by the environment. For example, we may compensate for our memories through repetition, writing, storytelling, etc. (Sutton, 2010; Suddendorf and Corballis, 1997).[[36]](#footnote-36) That a trait pervasively fails to perform its survival-value function may make little difference in the lifeway of the organism as a whole.

 At this point, I must leave off searching for ways in which the negative argument can be supported. I am not sure if there are any, since the argument faces serious obstacles. In any event, it remains the burden of those who wish to draw the revisionary conclusion to provide the relevant premises.

**Part 4. Conclusion**

 Much of the excitement around the simulationist program has been due to its revisionary character. It seems to many researchers that we are now discovering that memory is not what we had assumed–that the folk-psychological picture of memory has been fundamentally misguided. This suspicion finds its clearest expression in what this essay refers to as the negative project. On the basis of the negative project, simulationists have offered several hypotheses about what, if not remembering, episodic memory *is* for, why it emerged in the way it has, and how it works now. This essay argued that the relationship between the traditional storehouse view of memory and the simulationist approach is more complicated than has been acknowledged. In particular, this essay suggested that the storehouse view is not committed to idea that episodic memory only succeeds if it is fully accurate. How the accuracy of an episodic memory representation conditions whether retrieving it counts as episodic remembering is both complicated, and contextual. This makes it hard to draw conclusions about the frequency of remembering, and episodic memory function from false memory research directly. While the memory metaphors of Aristotle and Augustine are obviously far-fetched—there are indeed no waxen blocks, palaces, or caverns in the mind—we are not yet forced to abandon the old idea that memory is for remembering.

References

– Aristotle (2014) “Chapter Two: De Memoria et Reminiscentia: Text and Translation” In *Aristotle on Memory and Recollection: Text, Translation, Interpretation, and Reception in Western Scholasticism*, by Bloch, D., Leiden, NL: *Brill Academic Publishers.*

– Atance, C. M., & O’Neill, D. K. (2001) “Episodic future thinking” *Trends in Cognitive Science*, 12, 533-539.

– Augustine (1991) *Oxford Wrold Classics: Saint Augustine Confessions*, Trans. Henry Chadwick, Oxford, UK: Oxford University Press.

– Bartlett, F.C. (1932) *Remembering.* Cambridge, UK: Cambridge University Press.

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

– Bigelow, R. Pargetter, R. (1987) “Functions” *Journal of Philosophy* 84: 181-197.

– Boorse, C. (1976) “Wright on Functions”*The Philosophical Review* 85(1): 70-86.

– Boyer, P. (2009) “What are memories for? Functions of recall in cognition and culture,” in *Memory in Mind and Culture*, ed. P Boyer and J.V. Wertsch, Cambridge: Cambridge University Press, 1-26.

– Busby, J., Suddendorf, T. (2005) “Recalling yesterday and predicting tomorrow” *Cognitive Development*, 20, 362-372.

– Cheng, S., and Werning, M. (2016) “What Is Episodic Memory If It Is a Natural Kind?” *Synthese*, 193(5): 1345–85.

– Cummins, R. (1975) “Functional Analysis”. *The Journal of Philosophy*. 72(2), 741-765.

– Cummins, R. (2000) “’How does it work’ versus ‘what are the laws?’: Two conception of psychological explanation”. In F.Keil and Robert A. Wilson (eds.), *Explanation and Cognition*. Cambridge MA: MIT Press, 117-145.

– D’Argembeau, A. and Van der Linden, M. (2004) “Phenomenal characteristics associated with projecting oneself back into the past and forward into the future: infouence of valence and temporal distance. *Consciousness and Cognition*, 13: 844-858.

– De Brigard, F. (2014a) “Is Memory for Remembering? Recollection as a Form of Episodic Hypothetical Thinking”. *Synthese*, 191, 1-31.

 – (2014b) “The Nature of Memory Traces” *Philosophy Compass.* 9/6: 402-404.

– Deese, J. (1959) On the prediction of occurrence of particular verbal intrustions in immediate recall. *Journal of Experimental Psychology*, 58, 17-22.

– Garson, J. (2017) “How to Be a Function Pluralist” *The British Journal for the Philosophy of Science*, 69(4): 1101-1122.

– Godrey-Smith, P. (1993) “Functions: Consensus Without Unity” *Pacific Philosophical Quarterly* Vol.74 196-208.

– Godrey-Smith, P. (1994) “A Modern History Theory of Functions” *Noûs* (28) 344-362.

– Gould, S.J. (1980) “The Panda’s Thumb: More Reflections in Natural History,” New York, New York: W.W. Norton & Company, Inc.

– Gould, S.J., Lewontin, R.C. (1979) “The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme” *Proceedings of the Royal Society of London*. *Series B, Biological Sciences,* 205(1161): 581-598.

– Hassabis, D. Kumaran, D., Vann, D.S. et al. (2007) “Patients with hippocampal amnesia cannot imagine new experiences” *Proceedings of the National Academy of Sciences of the United States of America*,1014: 1726-1731.

– Hassabis, D., and Maguire, E.A. (2007) “Deconstructiong episodic memory with construction” *TRENDS in Cognitive Science*, 11(7): 299-306.

– Head, H. (1920) *Studies In Neurology*, London, UK: Oxford University Press.

– Hoerl and McCormack (2016) “Making Decisions about the Future: Regret and the Cognitive Function of Episodic Memory” In *Seeing the future: Theoretical perspectives on future-oriented mental time travel.* Michaelian, K., Klein, S., and Szpunar, K. Eds. Oxford UK: Oxford University Press. 241-266.

– Klein, S.B.,Loftus J, and Kihlstron J.F. (2002) “Memory and tempral experience: the effects of episodic memory loss on an amnesic patient’s ability to remember the past and imagine the future” *Social Cognition* (20) 353-379.

– Koriat, A. and Goldsmith, M. (1996) “Memory metaphors and the real-life/laboratory controversy: Correspondence versus storehouse conceptions of memory” *Behabioral and Bran Sciences* 19: 167-228.

– Lewontin, R.C. (1978) “Adaptation” *Scientific American*, 239(3) : 212-230.

– Locke, J. (1690/1999) *An Essay Concerning Human Understanding*. University Park, PA: Pennsylvania State University Press.

– Lloyd, E. “Adaptationism and the Logic of Research Questions: How to Think Clearly About Evolutionary Causes” *Biological Theory* 10:343-362.

– Loftus, E., Palmer, J. (1974) “Reconstruction of Automobile Destruction: An Example of the Interaction Between Language and Memory” *Journal of Verbal Learning and Verbal Behavior*, 13, 585-589.

– Loftus, E., Pickrell, J. (1995) “The formation of false memories.” *Psychiatric Annals*, 25: 720-725.

– Martin, C.B., and Deutscher, M. (1966) “Remembering,” *The Philosophical Review*, 75(2): 161-196.

– Michaelian, K. (2011) “Generative Memory” *Philosophical Psychology*, 24(3): 323-342.

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

– Millikan, R.G. (1984) *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, MA: MIT Press.

 – (1989) “In Defense of Proper Functions” *Philosophy of Science* 56, (2), 288-302.

– Neander, K. (1991) “Functions as Selected Effects: The Conceptual Analyst’s Defense” *Philosophy of Science*, 58(2): 168-184.

– Okuda, J., Fujii, T., Ohtake, H., Tsukiura, T., Tanji, K., Suzuki, K., et al. (2003) “Thinking of the future and the past: The oles of the frontal pole and the medial temporal lobes. *NeuroImage*, 19: 1369-1380.

– Plato (1997) “Theatetus” Trans. Levett, M.J., and Burnyeat, M. in *Plato: Complete Works*, Cooper, J.M. Ed. Indianapolis, IN: Hackett Publishing Company Inc.

– Quillian, M.R. (1966) *Semantic Memory*, PhD Dissertation, Carnegie Institute of Technology, In *Semantic Information Processing* Minsky, M Ed., Cambridge, MA: 227-270.

– Roediger, H.L.III and McDermott, K.B. (1995) “Creating False Memories: Remembering Words Not Presented in Lists” *Journal of Experimental Psychology*, 21, 803-814.

– Robins, S.K. (2016a) “Representing the past: memory traces and the causal theory of memory” *Pilosophical Studies*, 173: 2993-3013.

 – (2016b) “Misremembering” *Philosophical Psychology*, 29(3):432-447.

 – (2018) “Mnemonic Confabulation” *Topoi,* http://doi.org/10.1007/s11245-018-9613-x

– Schacter, D.L., Addis, D.R. (2007) “The cognitive neuroscience of constructive memory: remembering the past and imagining the future” *Philosophical Transactions of The Royal Society of Britain*, 362: 773-786.

– Schacter, D.L. (2012) “Adaptive constructive processes and the future of memory” *American Psychologist* 67(8): 603-613.

– Szpunar, K. K., Watson, J. M., & McDermott, K. B. (2007). “Neural substrates of envisioning the future,” *Proceedings of the National*

 *Academy of Sciences of the United States of America*, 104: 642–647.

– Suddendorf, T., & Corballis, M.C. Addis, (1997) “Mental Time Travel and the Evolution of the Human Mind” *Genetic, Social, and General Psychology Monographs*, 123(2), 133-167.

– (2007) “The evolution of foresight: What is mental time travel and is it unique to humans?” *Behavioral and Brain Sciences,* 32, 299-313*.*

– Sutton, J. (2010) “Exograms and Interdisciplinarity: History, the Extended Mind, and the Civilizing Process” In *The Extended Mind* Menary, R.

Ed. Cambridge, MA: MIT Press, 189-225.

– Tingerbergen, N. (1963/2005) “On aims and methods of Ethology” reprinted in *Animal Biology,* 55(4): 297-321.

– Tulving, E. (1972) “Episodic and semantic memory” in *Organization of Memory*, Tulving, E. Donaldson, W. Eds. N.Y.,N.Y.: Academic Press, Inc., 381-403.

 – (1983) *Elements of Episodic Memory*, Oxford, UK: Oxford University Press.

– (1985) “Memory and consciousness” *Canadian Psychologist*, 26: 1-12.

– (1993) “What Is Episodic Memory” *Current Directions in Psychological Science*, 2(3): 67-70.

– (2001) “Origin of autonoesis in episodic memory” In *The Nature of Rememebering: Essays in Honor of Robert G. Crowder*, Roediger, H.L. III, Nairne, J.S., Neath, I. and Surprenant, A.M. Washington DC: American Psychological Association.

– Wright, L. (1973) “Functions” *The Philosophical Review*, 82(2), 139-168.

1. Unfortunately, how to characterize episodic memory is controversial. Roughly, episodic memories are those which target particular experiences from one’s personal past (Cf. Robins, 2018). This suffices for a working definition, but researchers ultimately want an account that can distinguish episodic memory from “semantic” or merely factual memory. The field is divided on how to provide a more precise characterization. (See part two.) [↑](#footnote-ref-1)
2. For discussion of the duality of memory traces, physical and intentional, in Aristotle’s *De Memoria et Remiscentia*, see De Brigard (2014b, p. 402-403). [↑](#footnote-ref-2)
3. Aristotle, for example, explains memory problems by reference to physical problems of imprinting and storage. “[T]hose who are in much movement because of an affection or because of age do not come to have memory, as though the movement produced by sensation and the seal were impinged on running water, while others do not receive the impression because of damage in that which is receiving the affection–similar to the damage of old walls in buildings–and because of the hardness in it. Therefore, both the very young and old people have weak memories: for the former are fluctuating because of growth, the latter because of decay. Similarly neither those people that are too quick-witted nor those that are too slow-witted seem to possess a good memory: the former are moister than what is needed, the latter are harder; thus the image does not remain in the soul of the former, while it does not make real contact with the latter,” (450b1-11). Aristotle’s precise theory is irrelevant here. What I draw attention to is the general strategy of explaining memory errors by reference to physical and causal parameters. [↑](#footnote-ref-3)
4. The terms “sameness” and “identity” are ambiguous between two sense: qualitative and quantitative. If two items, *x* and *y* are qualitatively identical, then they share *some* properties in common. If *x* and *y* are quantitatively identical, then they are *completely* qualitatively identical, share *all* their properties in common, and hence anything true of *x* is true of *y*. For simplicity, the discussion above is couched in terms of quantitative or “numerical” identity. This should not indicate that comparable remarks can’t be made in terms of qualitative identity. [↑](#footnote-ref-4)
5. Certain passages of Locke suggest a similar view. Memory, for example, is described as “the power to revive again in our minds those ideas which, after imprinting, have disappeared or have been as it were laid out of sight…This is memory, which is as it were the storehouse of our ideas…this laying up of our ideas in the repository of the memory signifies no more but this – that the mind has a power in many cases to revive perceptions which it has once had, with the additional perception annexed to them, that it has had them before. And in this sense it is that our ideas are said to be in our memories, when indeed they are actually nowhere; – but only there is an ability in the mind when it will to revive them again, and as it were paint them anew on itself, though some with more, some with less difficulty; some more lively, and others more obscurely,” (Locke, p. 132-133). (For discussion, see De Brigard, 2014b). [↑](#footnote-ref-5)
6. Note, also, that perfect preservation of content from the representation deposited to the representation retrieved is not even sufficient for remembering, since the content of the representation deposited also has to have been accurate to begin with. [↑](#footnote-ref-6)
7. Perhaps, whether a subject counts as remembering *semantically* also exhibits a degree of context-sensitivity. For example, suppose Jones recalls (semantically) that, “Cape Verde is an island nation in the Atlantic Ocean,” but is unsure whether it is located in the Caribbean Sea or off the west coast of Africa. I take it that in some contexts, Jones may count as semantically remembering where Cape Verde is located, and in other contexts not, depending on whether the degree of precision that is relevant in the context requires distinguishing between the Caribbean and the mid-Atlantic. [↑](#footnote-ref-7)
8. A different way of describing the relevant relation is to say that, in remembering, the contents of the representations deposited/retrieved should be *determinate* properties of the same contextually-relevant *determinable* property. This suggestion also allows for the *tolerance* of remembering. As determinate properties, the contents of the representations deposited/retrieved may differ, so long as they are determinates of the same relevant determinable. [↑](#footnote-ref-8)
9. De Brigard (2014b), for example, contends that “Aristotle’s view – i.e., that memory representations are stored and causally derived copies of previous perceptions – is essentially identical to what appears to have been the received philosophical view on memory traces up until the early 1980s [i.e. the causal theory]…[i]nitially resurrected by Martin and Deutscher (1966),” (p. 403). [↑](#footnote-ref-9)
10. Martin and Deutscher’s account is the canonical causal theory of remembering. On their analysis, a subject remembers something exactly when:

1. “Within certain limits of accuracy he represents that past thing.

2. If the thing was ‘public,’ then he observed what he now represents. If this thing was ‘private,’ then it was his.

3. (a) His past experience of the thing was operative in producing a state or successive states in him finally operative in producing his representation,” (p. 166).

(b) “In those cases where prompting is operative for the representation, his past experience of the thing represented is operative in producing the state (or the successive set of states) in him which is finally operative in producing the representation, *in* the circumstances in which he is prompted,” (p. 185).

(c) “The state or set if states produced by the past experience must constitute a structural analogue of the thing remembered, to the extent to which he can accurately represent the thing,” (p. 191).

Efforts have been made by recent causal theorists to get clearer on how the physical trace(s) left by the relevant experience contributes to the production of the occurrent representation, since Martin and Deutscher’s structural analogue condition (c) is often considered untenable in light of current science. There are difficulties pressing connectionist representations into service of the causal theory (Cf. Robins, 2016a). [↑](#footnote-ref-10)
11. Ebbinghaus discovered the forgetting curve by memorizing lists of nonsense syllables (“CVC-trigrams”) and counting how many items he could retrieve at intervals. [↑](#footnote-ref-11)
12. For discussion of the relevant history in cognitive neuroscience, see, e.g., Squire, 2009; Squire, and Wixted, 2011. [↑](#footnote-ref-12)
13. Tulving cites Quillian (1966) as the first use of the term ‘semantic.’ [↑](#footnote-ref-13)
14. On one hand, “every ‘item’ in episodic memory represents information about the experienced occurrence of an episode or event,” (p. 387). On the other, “[i]f a person possesses some semantic memory information … he need not possess any mnemonic information about the episode of such learning,” (p. 389). [↑](#footnote-ref-14)
15. Tulving reports that K.C.’s “procedural and semantic memory systems are relatively unimpaired whereas his episodic system is severely damaged,” (1985, p. 5). [↑](#footnote-ref-15)
16. As Tulving describes it, “autonoetic consciousness, correlated with episodic memory…is necessary for the remembering of personally experienced events. When a person remembers such an event, he is aware of the event as a veridical part of his own existence,” (1985, p. 3). [↑](#footnote-ref-16)
17. For example, while Suddendorf and Corballis (2007), and Michaelian (2016) explicitly commit to the autonoetic criterion, Schacter and Addis (2007), and De Brigard (2014) do not commit to it. [↑](#footnote-ref-17)
18. Sometimes this argumentative strategy is more explicit than others. De Brigard expressly characterizes his methodology as inference to the best explanation. “[I]t is important to highlight that in attributing such a function (i.e., episodic hypothetical thinking) to this larger system I am merely making an inference to the best explanation,” (2014, p. 180). Moreover, that remembering and simulating are regarded as incompatible alternatives seems evident from quotes like the following. “According to its mechanistic role, remembering does not seem to be what remembering is for, as remembering appears to be a subroutine of a more complex operation [i.e. episodic hypothetical thought],” (Ibid., 168). Suddendorf and Corballis (2007) betray similar commitments when they write, “[t]he fact that episodic memory is fragmentary and fragile suggests that its adaptiveness may derive less from its role as an accurate record of personal history than from providing a ‘vocabulary’ from which to construct planned future events (and perhaps to embellish events of the past),” (p. 302-303). [↑](#footnote-ref-18)
19. A similar argument is revived by Robins’ (2016b) against the “archival” view of memory. [↑](#footnote-ref-19)
20. More specifically, Bartlett explains that, “when a subject is being asked to remember, very often the first thing that emerges is something of the nature of attitude. The recall is then a construction, made largely on the basis of this attitude, and its general effect is that of a justification of the attitude,” (207). Remembering, in other words, “is an imaginative reconstruction or construction, built out of the relation of our attitude towards a whole active mass of organized past reactions or experience… It is thus hardly ever really exact, even in the most rudimentary cases of rote recapitulation, and it is not at all important that it should be so,” (213). [↑](#footnote-ref-20)
21. The following is a study list from Roediger and McDermott (1995) that is focused around the lure of ‘sweet’. Sour, candy, sugar, bitter, good, taste, tooth, nice, honey, soda, chocolate, heart, cake, tart, pie (p. 814). [↑](#footnote-ref-21)
22. Roediger and McDermott strongly emphasize this point, (1995, p. 812). [↑](#footnote-ref-22)
23. Simulationists sometimes point to additional false memory effects. They are assigned the same argumentative work: to bolster the idea that episodic memory is often distorted. Here are some. The *boundary extension effect* occurs when a larger portion of a spatial scene is represented than had been perceived. The *telescoping effect* occurs when remote events are represented as more recent, and recent events are represented as more remote in time. *Observer perspective memory* occurs when a subject represents themselves in an event, as if seeing themselves from a third-person perspective. For discussion, see De Brigard (2014a, p. 159-162). [↑](#footnote-ref-23)
24. It is important to note that some simulationists may be more committed tote negative project than others. Schacter and Addis, for example, appear open to the possibility that remembering is *one* of episodic memory’s functions. [↑](#footnote-ref-24)
25. Episodic future simulation involved activity in frontal areas not implicated in episodic memory. [↑](#footnote-ref-25)
26. According to De Brigard, “[t]his core brain network involves the hippocampus, the posterior cingulate/retrosplenial cortex, the inferior parietal lobe, the medial prefrontal cortex, and the lateral temporal cortex,” (p. 174). [↑](#footnote-ref-26)
27. If I read them correctly, Schacter and Addis and De Brigard are in this camp. [↑](#footnote-ref-27)
28. For an early discussion of this idea, often called “function pluralism,” see Godfrey-Smith (1993). For more recent discussion of variations of the view, see Garson (2017). [↑](#footnote-ref-28)
29. This is Cummins’ official recipe: “*X* functions as a *φ* in *s* (or: the function of *x* in *s* is to *φ*) relative to an analytical account *A* of *s*’s capacity to *ψ* just in case *x* is capable of *φ*-ing in *s* and *A* appropriately and adequately accounts for *s*’s capacity to *ψ* by, in part, appealing to the capacity of *x* to *φ* in *s*,” (1975: 726). [↑](#footnote-ref-29)
30. There are important differences between the etiological accounts cited above. The discussion in the main text is modeled on Millikan’s approach. Her theory requires that only devices that are members of a “reproductive lineage,” or are produced by members of a reproductive lineage have functions. This restriction was included to avoid the counter-examples Boorse (1976) advanced against Wright’s account. Millikan’s definition is as follows. “[F]or an item *A* to have a function *F* as a “proper function”, it is necessary (and close to sufficient) that one of these two conditions should hold. (1) *A* originated as a ‘reproduction’ (To give one example, as a copy, or a copy of a copy) of some prior item or items that, *due* in part to possession of the properties reproduced, have actually performed F in the past, and *A* exists because (causally historically because) of this or these performances. (2) *A* originated as the product of some prior device that, given its circumstances, had performance of *F* as a proper function and that, under those circumstances, normally causes *F* to be performed by *means* of producing an item like *A*,” (1989: 288). On Millikan’s terminology, devices satisfying (1) have “proper functions,” while items satisfying (2) have “derived proper functions.” Organs, for example, have derived proper functions because they are not produced directly from past organs, but are produced from genes. [↑](#footnote-ref-30)
31. There is a serious epistemic obstacle to ascribing survival-value functions that are forward-looking. These are functional ascriptions based on the propensity for survival a device *currently* confers on its possessor. Traits only confer a propensity for survival on their possessors relative to a range of alternatives, and circumstances of selection. The problem is that it is obscure how to identify what the relevant ranges of alternative traits and circumstances are. For a thorough discussion of the problem, see Godfrey-Smith (1994, p. 352-353). [↑](#footnote-ref-31)
32. For a second example, De Brigard’s account is explicitly addressed to episodic memory’s “mechanistic role function” (Craver, 2001). Craver’s theory is a specification of Cummins’ causal-role account that is intended to make precise how systems performing causal role functions are organized in space and time. Yet, De Brigard argues for the negative project using considerations that seem appropriate to ascribe an etiological rather than a causal/mechanistic-role conception of function. For example, De Brigard asks, “why would we have a cognitive system that malfunctions so constantly and so systematically?” Questions about why we have certain systems seem to be etiological questions about why the system is here. But systems can perform causal-role functions that are not their etiological functions. So inferences from etiological grounds to causal-role conclusions need to be drawn with care. [↑](#footnote-ref-32)
33. Indeed, the canonical storehouse metaphors were proposed before Darwin developed the theory of evolution. [↑](#footnote-ref-33)
34. An example can be found in Lewontin (1978). “Every trait is involved in a variety of functions, and yet one would not want to say that the character is an adaptation for all of them. The green turtle *Chelonia mydas* is a large marine turtle of the tropical Pacific. Once a year the females drag themselves up the beach with their front flippers to the dry sand above the high water mark. There they spend many hours laboriously digging a deep hole for their eggs, using their hind flippers as trowels. No one who has watched this process would describe the turtle’s flippers as adaptations for land locomotion and digging; the animals move on land and dig with their flippers because nothing better is available,” (218). Here, it is plausible to say that while the green turtle’s hind flippers have the etiological function of supporting locomotion in water, they have a further, causal/mechanistic-role function of serving as trowels. [↑](#footnote-ref-34)
35. While considering principle R2, we have been bracketing the assumption that episodic memory is an adaptation. In ascribing an etiological function to episodic memory, this claim would need to be argued not assumed (Gould and Lewontin, 1979; Lloyd, 2015). For one thing, there would need to be evidence that episodic memory is itself heritable: a discrete genetic package or unit that natural selection could act upon. [↑](#footnote-ref-35)
36. Thanks to an anonymous reviewer for bringing the discussion of this point in Suddendorf and Corballis, 1997 to my attention. [↑](#footnote-ref-36)