THE EPISTEMOLOGY OF MEMORY BELIEFS

BY MARY SALVAGGIO

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Philosophy

Written under the direction of

Alvin Goldman and Brian McLaughlin

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2018

ABSTRACT OF THE DISSERTATION

The Epistemology of Memory Beliefs

by Mary Salvaggio

Dissertation Director: Alvin Goldman and Brian McLaughlin

Most of our beliefs are memory beliefs. It is rather surprising, then, that the epistemology of memory has been relatively neglected until recently. This is because memory beliefs have been traditionally understood as stored information. According to the traditional view, beliefs are stored in our memory similar to the way books are stored in a library, and just as books can get damaged, lost, or misplaced in the library, beliefs can be forgotten. The current consensus view of memory in psychology contradicts the traditional philosophical view. Memory is a constructive process; it is not mere storage and retrieval.

The dissertation addresses three major questions that arise in the philosophy of memory: What is memory? How are memory beliefs justified? And can forgetful agents be rational? Chapter 2 answers the first question. I defend a capacity account that claims that memory is a neurocognitive capacity to encode, store, and retrieve information. Unlike traditional philosophical accounts, the capacity account

is compatible with both current psychology and folk notions of memory. In Chapter 3, I argue for a process reliabilist view of the justification of memory beliefs. My view is distinctive in offering separate treatments of two importantly different types of processes that can generate memory beliefs, reproductive and reconstructive processes. Chapter 4 explores the final question. I make a previously unappreciated distinction between two types of forgetting: in addition to forgetting their evidence, an agent might also forget their conditional credences. I argue that thinking carefully about the rational status of forgetting conditional credences can improve our current theories of rationality in the same way that forgetting evidence has.

Acknowledgements

In writing this dissertation, I have received advice and encouragement from many. First of all, I thank my dissertation director, Alvin Goldman. From my first paper on memory in his proseminar until today, he has been there for me through it all. I also owe much to my committee members, Brian McLaughlin, Susanna Schellenberg, and Jack Lyons for their invaluable suggestions on drafts.

Outside of the committee, Preston Greene provided extensive comments on previous drafts of the dissertation. Bob Beddor, Will Fleisher, Matthew Frise, Simon Goldstein, Neil Mehta, Sara Robbins, Nico Silins, and Katherine Ward also provided useful feedback. Audiences at the Issues in Philosophy of Memory Conference, the National University of Singapore, the Bled Philosophy Conference, the Australasian Association of Philosophy Conference, the Philosophers' Cocoon Philosophy Conference, and Rutgers Philosophy and Cognitive Science graduate students provided helpful comments and criticisms as well.

A portion of this dissertation has been previously published elsewhere. Chapter 3 is slightly modified from Salvaggio (2018) which appeared in *Philosophical Studies*.

Dedication

To Preston, this wouldn't have been possible without your love and support.

Table of Contents

Chapter 1

Introduction

Most of our beliefs at any given time are the result of memory. Jenny believes that Washington D.C. is the capital of the United States of America. Danny believes that he had chocolate cake for his 10th birthday. Amelia believes that coffee shops usually have tables and chairs. A psychological study participant believes that they met Bugs Bunny in Disneyland as a child. All of these are memory beliefs.

It is rather surprising, then, that the epistemology of memory has been relatively neglected until recently. This is partly because memory beliefs have traditionally been understood as stored information. The view holds that beliefs are stored in our memory similar to the way books are stored in a library. In contrast, the current consensus view of memory in psychology is much more complex. Long-term human memory is not merely a matter of retrieving stored information. Each stage of the memory process—encoding, storage, and retrieval—can be influenced by context, background beliefs, and their interactions with the other stages.

Most importantly, many of the beliefs we take to be stored are actually generated at the time of retrieval. Rather than mere passive storage, memory is often a creative and interactive process. This fact must be addressed in philosophical views of memory. We now know that memory is so much more interesting and complex than we thought, and this realization creates the need for more nuanced accounts of the epistemology of memory. Embarking on that project is the main work of this dissertation.

The dissertation addresses three major questions that arise in the philosophy of memory:

- What is memory?
- How are memory beliefs justified?
- Can forgetful agents be rational?

By addressing these three important questions, this dissertation offers a comprehensive account of the epistemology of memory beliefs. In making claims about what memory and forgetting are, I argue that there are important distinctions that have been overlooked. Making these distinctions allows me to provide unique accounts of the justification of memory beliefs and the rationality of forgetful agents.

First, in Chapter 2, I address the question of what human long-term memory is. This question is crucial to the epistemology of memory since, as I show in Chapter 3, epistemologists often have different phenomena in mind when they make claims about what justifies 'memory beliefs.' Making explicit the relevant category of beliefs at the outset allows for a cohesive account of the epistemology of memory in all its forms. In this chapter, I defend a capacity account which claims that memory is a neurocognitive capacity to encode, store, and retrieve information. Memory beliefs are the result of exercises of this capacity. While our understanding of what counts as encoding, storage, and retrieval will change over time with more psychological discoveries, the capacity account represents the core notion of what memory is, and especially so as it pertains to the epistemology of memory.

Armed with the capacity account, in Chapter 3 I turn to what makes memory beliefs justified. In this chapter, I argue for a process reliabilist view of the justification of memory beliefs. There are two importantly different types of processes that can

generate memory beliefs; reproductive and reconstructive processes. Reproductive memory mainly involves the retrieval of previously stored information. I argue that this form of memory preserves justification; these memory beliefs are justified as long as the belief was justified at the time of encoding and the memory process is conditionally reliable. Reconstructive memory, in contrast, generates justification inferentially: such memory beliefs are justified if the process used to generate them is conditionally reliable and all of the input beliefs are justified at the time of retrieval.

Finally, Chapter 4 explores the question of whether forgetful agents can be rational. The traditional Bayesian account of rationality requires perfect memory: once a piece of evidence has been conditionalized on, it can never be forgotten. However, forgetting in this way does not seem like a failure of rationality. There are several alternative accounts of conditionalization that attempt to solve this problem. These accounts all focus on just one kind of forgetful agent: agents who forget their evidence. I argue that a different kind of forgetful agent is equally important: agents who forget their prior conditional credences. The distinction between conditional credences and evidence corresponds, in humans, to reproductive and reconstructive memory. Our prior conditional credal states are mostly reproductive while unconditional credal states (prior evidence) tend to be reconstructive. I show that these different kinds of forgetting should be addressed separately. While an agent who forgets their past evidence may still be rational, the case is less clear for the agent who forgets their past conditional credences.

1.1 Chapter Summaries

The rest of the dissertation consists of three chapters. Each chapter is written as a paper that can be read in isolation from the others. To accomplish this, I sometimes repeat basic conclusions from previous chapters.

Chapter 2: A Capacity Account of Memory

In this chapter, I argue for a capacity account of memory, according to which memory is a neurocognitive capacity to encode, store, and retrieve information. Unlike its major competitors, phenomenal and causal accounts, the capacity account picks out an epistemically interesting class of memory beliefs while accommodating the constructive nature of memory. Phenomenal accounts classify memory as having a certain phenomenal character, but the mental processes generating the phenomenal feeling of pastness are separate from the processes that generate the content and have no direct bearing on the truth of the resulting beliefs. To fix this problem, the capacity account focuses instead on the processes that generate the contents of memories. Meanwhile, causal accounts require a causal connection between the subject's current representation and their original representation, but recent psychological research shows that memory is constructed rather than passively retrieved, and so such a causal connection does not exist. Overall, the capacity account is the only way to appreciate human memory for the rich, constructive process that it is.

Chapter 3: The Justification of Reconstructive and Reproductive Memory Beliefs

Preservationism is a dominant account of the justification of beliefs formed on the basis of memory. According to preservationism, a memory belief is justified only if that belief was justified when it was initially held. However, we now know that much (if not most) of what we remember is not explicitly stored, but instead reconstructed when we attempt to recall it. Since reconstructive memory beliefs may not have been continuously held by the agent, or never held before at all, a purely preservationist account of memory does not allow for justified reconstructed memory beliefs. In this chapter, I show how a process reliabilist account can maintain preservationism about reproductive memory beliefs while accommodating the justification of reconstructive memory beliefs. I argue that reconstructive memory is an inferential process and that therefore the beliefs it produces are justified in the same way that other inferential beliefs are justified. Accordingly, my process reliabilist account combines a preservationist account of reproductive memory with an inferential account of reconstructive memory. I end by defending this view against objections.

Chapter 4: Can Forgetful Agents be Rational?

According to standard Bayesian conditionalization, forgetting is irrational. In most cases, however, forgetting does not seem like a matter of rationality. Accordingly, many alternatives to conditionalization have been proposed which view forgetting as arational rather than irrational. Nevertheless, the existing literature has mostly overlooked a crucial distinction between two types of forgetting: in addition to forgetting their evidence, an agent might also forget their conditional credences. Is it

irrational to forget one's conditional credences or is this also merely arational? This chapter canvasses several of the proposed alternatives to conditionalization that view forgetting evidence as arational and explores what verdict they might give for agents who forget their conditional credences. While I agree that forgetting evidence is not irrational, I believe that the jury is still out when it comes to forgetting conditional credences. I argue that thinking carefully about the rational status of forgetting conditional credences can improve our current theories of rationality in the same way that forgetting evidence has.

Chapter 2

A Capacity Account of Memory

2.1 Introduction

A woman is looking at a childhood picture of her father. She thinks back fondly on the hot-air balloon ride they took together. She recalls her father comforting her when she was frightened at the takeoff. She remembers how it felt to look down on her hometown. This seems like a run-of-the-mill case of autobiographical memory, but none of it actually happened. She is a subject in a psychology experiment looking at a doctored photograph.

A man describes a scene from a childhood trip to the beach. He can tell you what he was wearing, the picnic lunch they ate—all the contextual details that a typical recollection would include. While he knows that the event happened to him, he reports feeling as if he was looking at a photo of someone else's vacation. After suffering head trauma in an automobile accident, he can retrieve facts about past events, but his experience lacks that special feeling of reliving the past.

Are these still memories? Many philosophers would say they are not.

Causal accounts of memory require a causal connection between the subject's current representation and their original representation of the event. The woman's beliefs about the balloon ride cannot be causally connected to a prior representation because there was no prior representation. According to causal accounts, the balloon

ride is a merely apparent memory—a constructed story—and not a true 'memory' at all.

According to phenomenal accounts, memories are mental states with a particular type of phenomenal character. This 'feeling of pastness,' or autonoetic consciousness, is what distinguishes memories from other mental states. Without this special conscious experience, the man does not remember his beach vacation despite being able to give detailed and accurate information about it.

In this chapter, I argue for a capacity account of memory, according to which memory is a neurocognitive capacity to encode, store, and retrieve information. According to this account, both cases involve genuine memory. Unlike its two major competitors, the capacity account picks out an epistemically interesting class of beliefs as memory beliefs, and it does so in a way that is consistent with contemporary psychology of memory. As I will show, the main problem with the phenomenal account is that using phenomenal criteria to demarcate memory results in a distinction without a truth-relevant difference. The main problem with the causal account is that most human long-term memories involve some construction, and the same constructive psychological mechanisms produce both 'merely apparent' and 'real' memories. In other words, human beings rarely rely entirely on stored information about the past; memory is a constructive process. The capacity account, meanwhile, allows us to appreciate human memory for the rich, constructive process that it is.

2.2 Three Conceptions of Memory

In this chapter, I will be arguing in favor of a capacity account of human memory against more narrow views. As we will see, these three conceptions are not extensionally equivalent. The phenomenal and causal accounts add more restrictive conditions to the permissive capacity account.

According to the *capacity account*, memory is a "neurocognitive capacity to encode, store, and retrieve information" (Tulving 2005). Remembering is an exercise of this capacity. Given this, the details of human memory (i.e., the nature of encoding, storage, and retrieval in human memory) are something to be discovered by psychologists and neuroscientists. For example, empirical research has discovered that long-term human memory involves active reconstruction, rather than mere retrieval.

On the *phenomenal account*, memories must exhibit a particular type of phenomenal character, such as that characteristic of autonoetic consciousness. This type of view would limit memory to only (some) episodic memories. It would be possible to include semantic memory as well if the kind of conscious awareness is not limited to autonoetic consciousness.[1]

The *causal account* requires memory to be caused in the right way by past experience. This causation is most often thought to go via the retrieval of information stored in a memory trace. Books store information in the form of printed words or images, computer hard disks store information as magnetic fields, and, according to this conception, humans store information as memories in our brains. Remembering is retrieving that stored information. This retrieval often, but not always, has a certain phenomenal character to it; however, according to the causal conception, what

[1]This second position is not held by anyone that I am aware of, I will consider it below purely for the sake of completeness.

really matters is the causal connection.

Finally, I must point out that my usage of “memory," “remember," and similar terms is a departure from ordinary English usage as well as their use in much of the current philosophical literature.[2]

S remembering that p is often thought to require one or more of the following:

- p
- S is justified in believing that p
- S knows that p

Most views require memories to be true. Epistemic accounts of memory typically build in requirements along the lines of justification or knowledge.[3] For example, Williamson (2000) views memory as a species of knowledge. The account of episodic memory in Cheng and Werning (2016) requires factivity and ‘knowledge-likeness.’ They consider false memories to be ‘improper episodic memory.’

If one subscribes to these requirements, then some things that I count as memory will be considered merely apparent memories. On this view, ‘real’ memories have contents that are true, believed, justified, and/or knowledge. Many philosophers who argue for this do so on linguistic grounds. I do not wish to dispute those arguments; I am interested here in exploring the nature of the cognitive phenomenon. As will become clear, I am interested in a psychological conception of memory. Optical

[2]Although, see Bernecker (2010) for an argument against the knowledge requirement.

[3]Not all epistemic accounts require justification or knowledge. Debus (2010) requires only that “in order for an experience to count as an instance of remembering, it is necessary that the relevant experience have epistemic relevance for the subject when judging about the past; that is, it is necessary that the subject be disposed to take the relevant experience into account when judging about the past.”

illusions and hallucinations are important subjects in the psychology of vision, even though “see" is also factive.

2.3 Capacity Account

In this section, I will describe a capacity account of memory and show how it conforms to our best current understanding of the psychology of memory. On the capacity account, memory is a mental capacity, like perception. Memory is a mental ability to encode, store, and retrieve information. When cognitive psychologists study memory, they are primarily using the capacity conception (Schacter 2007).

In order to understand this conception, we must explore what these human mental abilities for encoding, storage, and retrieval are like in more detail. Here I will briefly mention two main theories about the nature of memory: systems and component processes. According to a systems framework, memory consists of several systems that “can be distinguished in terms of the different kinds of information they process and the principles by which they operate” (Squire, 2004, 174). For example, there are thought to be systems for declarative and non-declarative memory[4] and within declarative memory itself are the episodic and semantic systems. However, the thing that makes them all memory systems as opposed to systems of some other type, is that they all encode, store, and retrieve information; albeit in different ways. Alternately, Cabeza and Moscovitch (2013) defend a components processing framework according to which cognitive acts such as remembering are carried out by ‘process-specific alliances’ (PSAs) which are “small group[s] of brain regions working together to achieve a cognitive process” (52) as opposed to fixed memory systems. According

[4]Whether the capacity account includes non-declarative memory depends on a resolution of the debate over whether the encoding, storage, and retrieval account applies to non-declarative memory.

to this view, we can understand memory by studying the regions involved and their interactions. On the capacity conception then, PSAs that perform the functions of encoding, storing, or retrieving information constitute our memory capacity.

Any view of the nature of the memory capacity will have to accommodate the fact that the brain regions involved in memory are also involved in other tasks. For example, brain regions involved in episodic memory are also involved in imagination, navigation, and episodic future thinking (Hassabis and Maguire 2007). Semantic memory regions also play a role in object recognition, social cognition, language, and scene construction (Binder and Desai 2011). Also, semantic memory and episodic memory can influence each other at both encoding and retrieval (Greenberg and Verfaellie 2010).

Simulationist and constructivist theories of episodic memory are types of capacity account. For example, according to Michaelian (2016) "S remembers an episode e just in case: S now has a representation R of e and R is produced by a properly functioning episodic construction system which aims to produce a representation of an episode belonging to S's personal past" (107). The episodic construction system, among other functions, retrieves relevant semantic and sensory information in order to construct a representation of the past episode. Thus, remembering on this account is an exercise of our memory capacity. According to Michaelian's account, the fact that it is the result of the system aiming to 'produce a representation of an episode belonging to S's personal past' makes it an instance of episodic remembering. Episodic remembering is one flavor of exercise of our memory capacity, but it is far from the only one. As De Brigard (2014) argues, the episodic construction system also constructs episodic hypothetical thoughts in the same way. According to the capacity account, these episodic hypothetical thoughts are also generated by exercising our memory capacity.

In the remainder of this section, I will describe several examples of episodic memory errors. Instances of false memory provide insight into the workings of episodic memory in the same way that illusions provide insight into the workings of perceptual systems.[5] These examples illuminate the constructive nature of episodic memory by showing how the construction can go wrong.

Eyewitness testimony is extremely susceptible to manipulation after the experience of the event. In Loftus (1975), subjects watched a film of a traffic accident and were asked questions about it. Depending on the question asked, people seemed to remember the accident differently. For example, subjects gave significantly higher estimates of speed when they were asked how fast the cars were going when they *smashed into* each other than when asked how fast they were going when they *hit* each other. When asked the *smashed* question, subjects were more likely to falsely report shattered glass on the scene when there was none. Additional studies are reviewed in Loftus (2005).

In Roediger and McDermott (1995), subjects were asked to memorize lists of several related words. The lists were designed to exclude a word that was strongly related to other words. For example, in lists including 'bed,' 'awake,' and 'rest,' the word 'sleep' was omitted. When asked to recall the list later, or to recognize items that were on the list, a large number of subjects erroneously included 'sleep.' These results have been replicated and expanded in a large number of follow-up studies (Gallo 2010).

In Brewer and Treyens (1981), subjects were brought into an office with various

[5]While I explore the epistemological implications of these cases in Chapter 3 of this dissertaion; here I merely want to mention that they are of particular epistemic interest. Memory errors ought to play a role in the epistemology of memory on par with the place of illusions and hallucinations in the epistemology of perception.

objects in it. Some were objects that one would expect to find in an office at the time, like a typewriter and chairs, and others were not—a skull, for example. Many subjects reported seeing objects, like books, which would typically be found in an office but were not present in this case. Many subjects remembered incongruous items like the skull as well. In a more recent study using photographs of naturalistic scenes, Hemmer and Steyvers (2009) further examine the influence of background knowledge on episodic memory.

Cases of boundary extension[6] involve subjects being asked to study a photograph. They are later tasked with drawing the photograph from memory. In the original study, 95% of subjects drew an image with more content than the original photograph, as if the photograph was taken with a wider angle. The drawings include reasonable extensions of the scene depicted in the photograph but outstrip its content.

In Mazzoni and Memon (2003), subjects either read about an event or were asked to imagine the event taking place. Subjects were more likely to report remembering a non-experienced event when they imagined it versus merely reading about it. Under one condition of the experiment, 30 percent of subjects formed a false memory for a long-discontinued medical procedure after imagining it.

These cases show that there is no guarantee that we only episodically remember things we have previously believed. For instance, it is presumed that the subjects in the office experiment did not believe that there were books while they were actually in the office. It is *possible* that some subjects did indeed believe that books were in the office during the experiment, even though they could not see them, but this is unlikely to explain every case. In at least some instances, the belief that books are in the office comes about later; it is more aptly described as *constructed* rather than

[6]Intraub and Richardson (1989), Spanò et al. (2017)

reconstructed.

Even where the belief was held at the time of encoding, it is still not necessarily the case that the information was stored during the time between when it was first formed and when it is later reconstructed, as shown in the word list studies. Presumably, some subjects reconstruct the belief that a word like “bed" was on the list, even if this was in fact believed at one point (e.g., when they were looking at the word) and then forgotten.

2.4 Phenomenal Conception

The phenomenal conception identifies memory with a certain kind of conscious experience. While I think that these experiences are interesting and important in their own right, I will argue that it would be a mistake to equate such experiences with memory itself.

I will focus on two problems for the phenomenal conception. First, the phenomenal conception groups together beliefs generated by different cognitive processes and separates beliefs generated by the same cognitive process. This grouping is made on the basis of truth-irrelevant factors. Secondly, even if we consider phenomenological differences as epistemically important in their own right, in many instances, it is almost impossible to tell the difference between the required kinds of consciousness and other types of experiences.

Consider the case of R. B. reported by Klein and Nichols (2012).[7] R. B. could describe past events in his life with the rich detail typical of episodic memory. However, he did not feel the normal autonoetic awareness of these events. Instead, while

[7]There is some controversy over the legitimacy of Klein and Nichols’ use of R.B. as a case study. For my purposes, it is an interesting thought experiment at the least.

he knew that these events happened to him, he reported feeling as if he was recalling the actions of another person. Despite being able to retrieve the relevant information, his memories did not feel like they belonged to him. According to Klein (2014), a proponent of the phenomenal conception, R.B. does not remember. Memory requires autonoetic consciousness, and R.B. does not have it.

Another interesting phenomenon to consider is déjà vécu (O'Connor et al. 2010). Déjà vécu is the feeling that a novel experience has occurred before. It is similar to déjà vu, except an experience of déjà vu has a strange quality to it and does not feel like an ordinary memory in the way a déjà vécu experience does. This case shows that the phenomenal experience is not, on its own, sufficient for memory.[8]

The problem I see here is that the phenomenal conception places beliefs produced by the same process in different categories while lumping together beliefs generated by very different processes. If R.B. regains autonoetic consciousness, then only the quality of his experiences would change, while the content and causal history of his memory beliefs would remain the same.[9] Déjà vécu beliefs are actually produced via perceptual systems, but, on a sufficiency version of the phenomenal account, they are classified with beliefs generated by a memory system.

If the categories picked out by the phenomenal conception told us something about the probability that the content is true, then perhaps we could understand the desire to create such a cross-process identification. However, the presence or absence of the phenomenology tells us nothing about how likely the content is to be true. In fact, the mental processes that generate the phenomenal feeling of pastness are separate

[8]I am not aware of anyone who defends a sufficiency version of the phenomenal conception. However, as will be discussed below, several views in the epistemology of memory take the presence of the phenomenology as sufficient for the *prima facie* justification of memory beliefs.

[9]Surprisingly, R. B. did, in fact, eventually recover his autonoetic consciousness.

from the processes that generate the content. Using phenomenal criteria to define memory thus results in a distinction without a truth-relevant difference. From the perspective of an epistemic externalist, the categories picked out by the phenomenal conception are epistemically useless.

Some epistemic internalists would reject this problem by arguing for the use of phenomenal factors in theories of justification. An internalist might hold that if you currently seem to remember that P, then you are currently justified in believing it. For example, Audi (1995, 37) claims:

> "If, for any proposition p that is not obviously false, (a) S has a memorial sense that p or (b) would (other things equal) have such a sense upon considering p, then S has prima facie memorial justification for believing p, of a degree sufficient to make it not unreasonable for a rational person in S's position to believe p."

Pollock (1974, 193) presents a similar view of the justification of memory beliefs. His view is that "'S recalls that P' is a prima facie reason for S to believe that it was true that P," where 'S recalls that P' just means that S is in a particular kind of phenomenal state.

Internalists who take this route are, however, faced with the following problem: how can we specify the kind of experience that makes something a memory? Psychologists tell us that there are two types of memory-feel experiences: autonoetic and noetic consciousness. The most distinctive kind of experience, autonoetic consciousness, has been described as "mental time travel,"[10] and is associated with episodic memory. Semantic memories tend to have a feeling of pastness about them, but this does not involve anything like the robust feeling of mental time travel. The feeling of familiarity that often accompanies semantic memories is called noetic awareness.

First, consider identifying memory beliefs on the basis of autonoetic awareness. As we have seen in the case of R.B., this awareness can be absent while the retrieval

[10]This phrase is attributed to Endel Tulving.

process is otherwise unaffected. Additionally, the same feeling is very similar to the conscious experiences related to other types of episodic thinking, especially the imagining of future events and counterfactual past events. One potential difference is that episodic memories tend to be more vivid and contain more sensory details (De Brigard and Giovanello 2012). However, it is a familiar problem that this kind of vivid mental imagery is not limited to the output of the brain's memory systems. As Plantinga (1993) notes, it is conceivable that a person's mental imagery is the same when they both remember and imagine a kind of vivid event, such as receiving a box of raspberries. Conversely, many episodic memories do not come with vivid mental imagery. In fact, some people claim to have no imagery whatsoever when they remember. Presumably, we would not want to exclude them from having memory beliefs.

Let us now turn to the other type of phenomenal experience, noetic awareness. It is also problematic for identifying memory beliefs because the feeling of familiarity can be confused with other kinds of experience. Noetic awareness is often so weak that there is little to distinguish it from plain old believing. Familiarity is not easily distinguishable from general processing fluency, which is the feeling of ease that can accompany processing that is less consciously effortful (Alter and Oppenheimer 2009). Previous experience with an item is just one way of increasing fluency, but items fluently processed for other reasons can also produce feelings of familiarity. For example, an item displayed in an easier to read font or in better focus than other items is more likely to be judged as familiar.

Additionally, the phenomenon of cryptomnesia occurs when people mistake a memory for a newly formed belief. This can lead to inadvertent plagiarism. One supposed example of this is discussed in Jung (1968). A passage from Nietzsche's

Thus Spake Zarathustra is almost a word for word reproduction of a story in a book published a half century before. Nietzche's sister confirmed to Jung that he had read the book as a child. Cryptoamnesia can also result in a kind of 'self-plagiarism.' Skinner (1983) writes "One of the most disheartening experiences of old age is discovering that a point you just made—so significant, so beautifully expressed—was made by you in something you published long ago." The absence of a feeling of familiarity is no guarantee that something is not a memory.

2.5 Causal Conception

On the causal conception, a memory belief is a proposition that has been stored in memory for later retrieval. While the causal conception enjoys some intuitive appeal,[11] ultimately it cannot account for what we know about how human long-term memory functions. As we saw in Section 3, the current consensus view in psychology is that much of human memory is reconstructed rather than simply retrieved from storage. Thus, the restriction to content that is explicitly stored is too limiting. In this section, I explore several causal accounts of memory and provide experimental cases where the construction goes beyond the level allowed for in the account.

Martin and Deutscher (1966) provide the classic causal account of memory, according to which memory requires a causal connection between an agent's current representation of something and their past experience of it. Causal accounts differ on the precise nature of the causal connection and the degree of similarity between the original experience and the memory representation. The important feature of Martin

[11] An intuitive feature of the view is its straightforward explanation of non-occurant memory. On the causal conception, a currently inaccessible belief can still be a memory belief (as long as it has not been completely forgotten). In this sense, you can be said to remember something even in cases where you fail to recall it, are not attempting to recall it at the moment, or, indeed, even if you are asleep or otherwise incapable of attempting to recall it right now.

and Deutscher's view, for our purposes, is that the causal connection involves *memory traces*. These traces are supposed to be structural analogs of the thing experienced and provide the causal link from a past experience to the current representation.[12]

Modern causal accounts agree that the causation involves memory traces of some sort. According to Bernecker (2009, 242), in order for a subject, S, to remember that p at $t2$ it must be the case that

> "S's representation at t2 that p is causally connected to S's representation at t1 that p* such that [i.] S's representation at t1 that p* and S's representation at t2 that p are connected by a persisting memory trace or a contiguous series of memory traces, [ii.] the memory trace is at least an inus ["insufficient but non-redundant factor of an unnecessary but sufficient" (143)] condition for S's representation at t2 that p. If the memory trace is an independently sufficient condition, it is not preempted by another independently sufficient condition, [and iii.] if S hadn't represented at t1 that p*, he wouldn't represent at t2 that p."

Where p* is a proposition that S represented at some past time (t_1) whose content is sufficiently simiar[13] to the remembered proposition, p. While the details of this account differ from that of Martin and Deutscher, the main idea of a causal connection to past experience via memory traces remains constant.

Memory traces can be used to distinguish genuine remembering from the effects of prompting and from relearning. Martin and Deutscher (1966, 186) give the case of accident-prone Kent, who has told a friend, Gray, about an accident Kent was recently involved in. Later, our unfortunate protagonist Kent develops amnesia about this period of time due to a head injury. Gray recounts the details of the original

[12]Setting aside the idea of a structural analog, this is broadly compatible with psychological definitions of memory traces. For example, Tulving (2007) describes a memory trace as "the neural change that accompanies a mental experience at one time (time 1) whose retention, modified or otherwise, allows the individual later, (at time 2) to have mental experiences of the kind that would not have been possible in the absence of the trace." See (Sutton and Windhorst, 2009) for an in-depth discussion of Martin and Deutscher's account of memory traces and modern psychological views.

[13]Sufficient similarity is spelled out in the following condition: "p and p* supervene on the same environmental conditions at t_1 or p is entailed by p* (where 'entailed' is understood along the lines of relevance logic.)" Bernecker (2009, 242) This content requirement is also in conflict with constructed memory, but I will only discuss the causal condition here.

accident to Kent who accepts them. Later, Kent forgets that he needed to be told about the accident by Gray, but remembers the details. Intuitively, Kent does not remember the accident. Kent's retelling of the accident is not caused by the memory traces of the original event (if they still exist), but by Gray's information.

However, prompting does not automatically preclude remembering. Take a case where someone, after a complete prompting of the events of a previously forgotten afternoon, claims that it is all coming back to him now and relates all the details back again. It is possible that this person really does come to remember after the prompting. We can assume that when 'it all comes back,' the prompting has activated the appropriate memory trace in some way.

Unfortunately, this explanation breaks down when confronted with cases of constructed memory. Compare accident-prone Kent to Kerri, a subject in a false memory experiment. Unlike Kent, Kerri is not suffering from amnesia; her memory is typical. Instead of being directly provided with the details of the accident, Kerri is asked a series of leading questions. Both Kent and Kerri can provide the detail that a sedan crashed into their car. Kent was told this information directly while Kerri was asked what color the sedan was that crashed into her car. According to the causal accounts of Martin and Deutscher and Bernecker, neither Kerri nor Kent remembers. Their current representation lacks the right sort of causal connection to past experience. They are able to state the detail about the sedan only because of the prompting they received.

As discussed in Section 3, many ordinary cases of remembering are similar to Kerri's in the relevant respects. Accurate memories that causal theorists would not wish to exclude are generated by the same kind of mental processes that result in the false memories. We know for sure that there can be no causal connection in the false

memory cases since no past experience exists; there is no reason to think that there is a causal connection in the similar cases of accurate memory. The pervasiveness of these constructed memories means that rejecting them as instances of remembering would significantly impoverish any resulting discussion.

This is not a problem that a causal account can solve with further refinements to the precise nature of the causal relation. The idea of a memory trace is integral to any causal account. Martin and Deutscher (1966, 189) describe the situation as follows:

> "Once we accept the causal model for memory we must also accept the existence of some sort of trace, or structural analogue of what was experienced. Even if someone could overcome the many difficulties of various kinds surrounding the idea of action at a distance, it could not be true to say that someone was remembering an event if his past experience of that event caused him, over a temporal gap, to recount it. There is an inevitable recourse to metaphors about the storage of our past experience involved in our idioms and thought about memory."

Since one cannot have a causal account without storage, memory that goes beyond stored content is incompatible with causal accounts.

This storage requirement rules out the possibility of false memory. As Martin and Deutscher (1966, 189) put it later in the same paragraph, "So long as we hold some sort of 'storage' or 'trace' account of memory, it follows that we can remember only what we have experienced, for it is in our experience of events that they 'enter' the storehouse." We have seen that in some cases, it is impossible for there to be any causal connection whatsoever to a past experience or representation since there was no past experience or representation to begin with. To capture the vastly constructive nature of human memory, we must reject the causal conception.

It is important to point out that this is only an argument that the causal conception cannot be the whole story about memory. I do not wish to deny that many memories are as the causal theorist describes; just that not all of them are. Even

though it might be incorrect to consider memory purely as a passive storehouse of information waiting to be accessed, some things must be stored even if we allow that not everything is.[14] The capacity conception of memory can account for the vastly constructive nature of human long-term memory. Remembering is not merely a matter of recalling previously stored information. Memory traces, stored background knowledge, and contextual details can all be recruited in the retrieval process. This can lead to representations that are natural extensions of stored content (as in boundary extension), can change details of stored episodes (as in the eyewitness testimony examples), or even generate entirely new content (as in the doctored photographs). Requiring strict correspondence with previously encoded and stored information would rule out a large and philosophically interesting portion of human long-term memories.

2.6 Conclusion

I have argued here for a capacity account of memory, according to which memory is a neurocognitive capacity to encode, store, and retrieve information. Unlike its two major competitors, the capacity account picks out an epistemically interesting class of beliefs as memory beliefs, and it does so in a way that is consistent with contemporary psychology of memory. The main problem with the phenomenal account is that using phenomenal criteria to demarcate memory results in a distinction without a truth-relevant difference. The main problem with the causal account is that most human long-term memories involve some construction, and the same constructive psychological mechanisms produce both 'merely apparent' and 'real' memories. In other words, human beings rarely rely entirely on stored information about the past;

[14] Additional reasons to hold on some aspects of the storage account can be found in Robins (2016) and Chapter 3 of this dissertation.

memory is a constructive process. The capacity account, meanwhile, allows us to appreciate human memory for the rich, constructive process that it is.

Chapter 3

The Justification of Reconstructive and Reproductive Memory Beliefs

3.1 Introduction

So many of our beliefs are memory beliefs. Right now, I believe that John was at the party last weekend, that eight times eight is sixty-four, that George Washington was the first president of the United States, that the movie last night made me sad, and that the substance I'm drinking is called 'coffee.' Am I justified in believing these things?[1] The answer depends on the way these beliefs are formed and maintained. Over the last century, psychologists have been advancing our understanding of how memory works, and some of the insights are surprising. We should ensure that our epistemology can accommodate the modern view of memory. In this chapter, I defend a theory of the justification of memory beliefs that is epistemologically satisfying while taking the psychological facts into account.

One of the most remarkable breakthroughs in psychology is the discovery that memory is reconstructive as well as reproductive. Our memories are not simple recordings of past experiences; they can be affected by our current context as well as background beliefs and other memories. Much of what we remember is not explicitly stored, but is instead constructed or reconstructed when we attempt to recall it. This

[1] Throughout, unless otherwise stated, I am interested in prima facie justification.

poses a problem for one of the dominant epistemological views of memory, preservationism. Preservationism is the view that memory cannot generate justification, but only preserve any original justification it had when it was formed.[2] Any beliefs reconstructed by memory cannot be justified according to this view. The preservationist has three choices: she can either bite the bullet and accept that most of our memory beliefs are unjustified, deny that reconstructed memories really are memories, or modify her account to allow reconstructed memories to be justified. I defend the third option. Reconstructive memories are generated via inferential processes, and so, I argue, are justified in the same way as other inferences.

In Section 2, I give a brief overview of the three traditional views of the justification of memory beliefs: foundationalism, preservationism, and inferentialism. I suggest that we want a view that avoids radical skepticism while also avoiding gullibility (being justified in too many beliefs) and rehearse arguments that preservationism satisfies these desiderata. However, I show that preservationism does not, on its own, provide a satisfactory account of the justification of reconstructed memories. The traditional preservationist picture is silent on cases where beliefs are reconstructed by memory. Section 3 examines the psychological evidence that many of our memories are, in fact, reconstructed. In Section 4, I present a view that supplements preservationism for reproductive memories with inferentialism for reconstructive memories. Finally, in Section 5, I defend this view against objections.

---

[2]Many preservationists claim that *knowledge* is preserved, but I am focused on the weaker claim about justification. See Lackey (2005) for compelling arguments against knowledge preservation.

3.1.1 Terminology

Before continuing, I must first clarify what I mean by 'memory belief.' Is a memory belief a belief that is caused in a particular way, a belief with a certain phenomenal feel, or both? It is easy to see that these accounts are not extensionally equivalent. Martin and Deutscher (1966, 167–168) describe a painter who paints a scene from his childhood but insists that the painting was a product of his imagination rather than his memory. Déjà vu is thought to be a feeling of remembering caused when current experience is similar to remembered experience (Cleary 2008). The imaginative painter has a belief that is caused in the right way, but it feels like something imagined rather than something remembered. In déjà vu, the belief has the right phenomenology, but it is caused by present rather than past experience.

As I am interested in applying the psychology of memory to its epistemology, I mean to be discussing beliefs that are caused in a particular way. By 'memory belief,' I refer to beliefs that result from an exercise of a particular mental faculty, memory, where memory is something like the "neurocognitive capacity to encode, store, and retrieve information" (Schacter and Tulving 1994). Not every belief resulting from an exercise of this capacity will have a 'memorial feel' and not every belief with such a phenomenology will be a result of the memory capacity. Accordingly, the painter counts as having a memory, but someone experiencing déjà vu doesn't. As we willl see, accepting this account of memory belief allows the philosopher to take advantage of the vast area of psychological work on memory.

It is also important for our purposes to allow for memories to be false. In English, 'remember' and its cognates require knowledge, or at least truth, for their application. If all memory is knowledge by definition, then the question of which of our memory beliefs are justified never arises. Instead, one must ask which of our apparent memory

beliefs are actual. I find the former question to be a more elegant statement of the issue. I will use 'remember' and 'memory belief' to refer to the psychological process and the beliefs resulting from that process, whether those beliefs are true or false; justified or unjustified. A reader who finds my choice of language offensive may feel free to insert 'apparent' in front of the offending phrases as necessary. In this chapter, when I write of 'merely apparent' memory beliefs, I am referring to beliefs that feel like memories, but were not produced via the cognitive faculty of memory.

3.2 Traditional Theories

Traditionally, there are three different views of how memory beliefs are prima facie justified; foundationalism, preservationism, and inferentialism.[3] Justification is a three place relation between a subject, S, a proposition, P, and a time, t.[4]

Foundationalism S is prima facie justified in believing P on the basis of memory at t if S (seems to) remember that P at t.[5]

Preservationism S is prima facie justified in believing P on the basis of memory at t if

1. S was prima facie justified in believing P at the time the belief was formed, $t_0 < t$, and

[3]As we will see, it is somewhat strained to call inferentialism an account of the justification of memory beliefs, per se. Rather, it is an application of a general view of justification to cases of memory beliefs.

[4]I am interested in doxastic, rather than propositional, justification. In what follows, I use "basis" to mean causal basis. In the definitions below, "believing P on the basis of memory at t" means that the subject's belief that P at t is a result of a memory process. A belief with a particular causal basis will have very different justificatory bases depending on which of the views below is accepted. For example, according to inferentialism, the justificatory basis for a memory belief is not memorial at all.

[5]Some externalist foundationalists will add a requirement that seeming to remember is a reliable process. See Plantinga (1993).

2. S's belief that P was stored in memory between t_0 and t.

Inferentialism S is prima facie justified in believing P on the basis of memory at t if S has non-memorial evidence that P is true.[6]

To illustrate these views, consider the following case. Jenny believes on the basis of memory that Washington D.C. is the capital of the United States. Is she prima facie justified in this belief?

Foundationalism is the most straightforward of the three traditional views. According to an internalist version of foundationalism, If Jenny seems to remember it, then she is justified. An externalist foundationalist will have an additional external requirement, such as requiring that remembering be a reliable process and/or that she actually remembers rather than merely seems to. Assuming that these additional requirements are fulfilled, Jenny is justified.

According to preservationism, we will need to know the history of Jenny's belief. How did she originally learn that Washington D.C. is the capital of the United States? If she first believed it justifiedly—for example, she learned it from a trustworthy geography teacher—then her current memory belief is also justified. However, if she originally believed it unjustifiedly, then her current belief is still unjustified as long as she has not gained any new evidence in the meantime.

According to inferentialism, the fact that Jenny seems to remember that Washington D.C. is the capital of the United States provides no (or negligible) justification on its own. The only way it can be justified is if she has evidence that makes her memory belief likely to be true. For example, she might justifiedly believe that her memory is generally reliable in these sorts of cases, or she might believe that the

[6]This explication of the view is neutral on what it takes for a subject to have evidence as well as what counts as evidence.

United States Capitol Building is in Washington D.C. and that capitol buildings are located in capital cities. However, if Jenny has no such justified beliefs, then her belief that Washington D.C. is the capital of the United States is not justified.

3.2.1 Problems with Traditional Views

There is much to say about these three views; however, I have space here for only a brief overview.

The main objection to foundationalism is that it gives us too many justified beliefs. If foundationalism were true, the mere fact that a belief is an apparent memory would be prima facie justification for that belief, regardless of its past justificatory status. We often believe things we are not justified in believing, and sometimes we later seem to remember these things. According to the foundationalist, recalling the belief grants it justification even if it had no justification before. For example, consider a 'forgotten carelessness' case. Suppose that you hear from a source you know is unreliable that Americans landed on the Moon in 1969. You believe it anyway, perhaps out of a sense of patriotism. Clearly, you are not justified in this belief when you form it.[7] Time passes, and you later recollect the fact. Are you justified in believing this now that you seem to remember it? According to the foundationalist, you are. Intuitively, you are not.[8]

Inferentialism faces either skepticism or circularity depending on how memory beliefs are supported by other evidence. There are two options here. First, you might

[7]This kind of objection is given by Senor (1993), Goldman (1999), and Huemer (2008).

[8]A committed foundationalist might bite the bullet here and accept the result that you are justified in these beliefs. For example, Schroer (2008) claims that you must accept it in order to take the foundationalist position seriously at all. At that point, the argument is at somewhat of a stalemate.

justify a memory belief with something like the following argument: (1) This belief is an apparent memory belief. (2) Apparent memory beliefs are highly likely to be true. Therefore, (3) this belief is highly likely to be true. It is hard to see how (2) will be justified without relying on memory.[9] Second, you might require additional justification for each memory belief, rather than an overall assessment of the reliability of memory. For example, suppose Jane remembers that every equilateral triangle is equiangular. She also remembers a proof of this fact. Surely Jane is justified in believing this. However, it is not clear that she is justified on an inferentialist view. If the premises are remembered, but not their justification, then Jane is not justified after all; the premises are also memory beliefs, and so they too require an inferentialist justification. On the other hand, if Jane remembers the entire proof from first principles, then she is justified. However, now it seems like memory has nothing to do with it. She is justified only if she can rehearse a proof of it from first principles, and if Jane can do that, the fact that she seems to remember it is beside the point.

Preservationism avoids the problems of foundationialism and inferentialism. In cases where the original belief was unjustified, that status is preserved when it is remembered. It also avoids skepticism by requiring past justification rather than independent evidence. Still, according to preservationism too few memory beliefs will be justified. Since memory can only preserve justification, beliefs will only be justified if they were actually believed in the past. If the subject has not held the belief at some earlier time and continued to hold it in memory, there is no original justification to preserve. In the following section, I will give an overview of some work

[9]Something like: "I have an apparent memory belief that, in the past, most of my apparent memory beliefs have turned out to be true."

in psychology that makes this problem especially pressing. Then I will propose an account that solves this problem.

3.3 Reproductive and Reconstructive Memory

Most contemporary psychologists agree that there are several memory systems.[10] Each memory system is a "neurocognitive capacity to encode, store, and retrieve information" Tulving (2005, 36). In what follows, I will focus on two memory systems: semantic and episodic.

3.3.1 Semantic Memory

A semantic memory is a memory of a fact without a memory of any specific experience. For example, you remember that George Washington was the first president of the United States. All you remember is the bare fact. You do not remember hearing it, or reading it, or any other particular experience. Experienced events can also be the subject of semantic memories. For example, you can remember that you had a chocolate cake for your birthday last year without being able to recall your birthday party.

3.3.2 Episodic Memory

Episodic memory incudes contextual information about experienced events, including how things looked, sounded, and smelled, as well as the emotions that were experienced. You can remember a delicious dinner in a way that you cannot remember

[10]The precise number of distinct memory systems, details about their nature, and their realization in the brain are all hotly contested. However, the details I rely on are relatively uncontroversial.

George Washington being the first president of the United States.

Psychologists have put forth several different kinds of evidence for this distinction. One of the most prevalent is dissociations, where subjects experience impairment in one system while the other is largely unaffected. These dissociations are usually found in patients with certain types of brain damage. For example, the patient K.C. suffered severe damage to his hippocampus and medial temporal lobes. He suffered from complete amnesia when it came to remembering events and situations, but his memory for facts, even those about his own life, was unimpaired. K.C. and patients with similar patterns of impairment suggest that episodic memory is separate from semantic memory (Tulving 2002). Damage to certain parts of the brain disrupt certain brain processes; since episodic memory is more effected than semantic memory when these processes are disrupted, these processes must be more involved in episodic memory performance than in that of semantic memory.[11]

Cubelli (2010) proposes a distinction between reproductive and reconstructive memory. Reproductive memory stores encoded information to be retrieved at a later time. Reconstructive memory uses stored information to construct a belief about a past experience.[12] The dominant view in psychology is that many episodic memories are reconstructive while semantic memories are reproductive. With reproductive memory, information retrieved is different from the information encoded by the memory system. Individual events are not stored with all details intact; instead, only a trace is stored. This trace, or a combination of traces, is then combined with general background beliefs and contextual information to reconstruct the episodic memory

[11] Another way of establishing that different brain areas are responsible is by using brain imaging. An examination of that evidence is beyond the scope of this chapter, but see Gabrieli (1998) for a review.

[12] Of course, one may not believe the output of a reconstructive memory process; however, I will focus only on beliefs here.

that is available to consciousness.[13] The content of a memory formed in this way may not have been encoded or stored before the retrieval occurred. This is in contrast to reproductive memory where information that is encoded is the same (or very similar) to the information retrieved.

Reproductive memories can be described using a storehouse metaphor. According to this conception, memory is like a library. Information is encoded and stored in order to be retrieved at a later time, like information written in a book is stored in a library to be read later. Reconstructive memories cannot be described with the storehouse conception. As will be discussed below, because of episodic memory's reconstructive nature, it is misleading to think in terms of beliefs and contents that are stored for later retrieval. Reconstructive memory is not like a camera that takes snapshots of events for us to look at later. Koriat (2007, 243) claims that reconstructive memory is "an intentional, goal-directed attempt to reconstruct a memorial representation from a variety of pieces of information that come to mind, negotiating between different considerations in attempting to arrive at a faithful account of previously encountered events." Instead of a library, reconstructive memory is more analogous to a detective, determining what most likely happened using the available clues.

While reconstructive memory can sometimes lead to errors, it also allows us to imagine the future, update our beliefs, and to be more creative (Schacter et al. 2011). Studying memory errors is similar to studying perceptual illusions in order to understand perception in normal cases. The kinds of errors that are made reflect the workings of the cognitive system that produces them. I will briefly summarize some of the most studied types of errors here.

---

[13]What kind of content traces have is controversial. I wish to remain neutral on this question. If they have belief-like content, then their justificatory status should be taken into account as well.

Eyewitness Testimony

Eyewitness testimony is extremely susceptible to manipulation after the experience of the event. In Loftus (1975), subjects watched a film of a traffic accident and were asked questions about it. Depending on the question asked, people seemed to remember the accident differently. For example, subjects gave significantly higher estimates of speed when they were asked how fast the cars were going when they *smashed into* each other than when asked how fast they were going when they *hit* each other. Additionally, when asked the *smashed* question, subjects were more likely to falsely report shattered glass on the scene when there was none. Additional studies are reviewed in Loftus (2005).

Word Lists

In Roediger and McDermott (1995), subjects were asked to memorize lists of several related words. The lists were designed to exclude a word that was strongly related to other words. For example, in lists including 'bed,' 'awake,' and 'rest,' the word 'sleep' was omitted. When asked to recall the list later, or to recognize items that were on the list, a large number of subjects erroneously included 'sleep.' These results have been replicated and expanded in a large number of follow-up studies (Gallo 2010).

Objects in a Scene

In Brewer and Treyens (1981), subjects were brought into an office with various objects in it. Some were objects that one would expect to find in an office at the time, like a typewriter and chairs, and others were not—a skull, for example. Many subjects reported seeing objects, like books, which would typically be found in an office but were not present in this case. Many subjects remembered incongruous items

like the skull as well. In a more recent study using photographs of naturalistic scenes, Hemmer and Steyvers (2009) further examine the influence background schemas on epsodic memory.

Imagination Inflation

In Mazzoni and Memon (2003), subjects either read about an event or were asked to imagine the event taking place. These events were either common and frequently occurring or ones that subjects could not have actually experienced—a long-discontinued medical procedure, for example. For both types of event, subjects were more likely to report remembering a non-experienced event when they imagined it versus merely reading about it. Under one condition of the experiment, thirty percent of subjects formed a false memory for the never-performed procedure after imagining it.

Since episodic memory is reconstructive rather than reproductive, there is no guarantee that we only episodically remember things we have previously believed. The false memory studies mentioned above are all examples of this. In *Objects in a Scene*, for instance, it is presumed that the subjects did not believe that there were books while they were actually in the office. Perhaps some subjects did indeed believe that books were in the office during the experiment, even though they could not see them, but this is unlikely to explain every case. In at least some instances, the belief that books are in the office comes about later; it is more aptly described as *constructed* rather than reconstructed. In cases where the belief was held at the time of encoding, it is still not necessarily the case that the belief exists during the time between when it was first formed and when it is later reconstructed, as shown in *Word Lists*. Presumably, some subjects reconstruct the belief that a word like "bed"

was on the list, even if this was in fact believed at one point (e.g., when they were looking at the word) and then forgotten. We need an account of the justification of memory beliefs that can accommodate these facts. In the remainder of this chapter, I will present such an account.

3.4 Preservationism and Inferentialism

In this section, I develop a process reliabilist account of memory. The basic idea behind process reliabilism is that the justification of a belief should depend on the reliability of the process or processes that caused the belief (Goldman 2008). Processes have inputs and outputs. Process reliabilism is interested in processes that have a belief as an output. Reliabilists typically distinguish two main types of processes depending on what kinds of inputs they take. A belief-independent process has no beliefs among its inputs. A belief-dependent process is one in which some of the inputs are beliefs. For a belief resulting from a belief-dependent process to be justified, all the input beliefs must themselves be justified and the process must be conditionally reliable. Conditional reliability is the reliabiliy of the process on the condition that all the input beliefs are true.

Memory is usually considered to be a prototypical belief-dependent process. As explained above, reproductive memory and reconstructive memory are actually two different kinds of processes. Reconstructive memory is clearly a case of a belief-dependent process. Reproductive memory is also a belief-dependent process, but in a different way—the input belief is that very belief from an earlier time. I claim that long-term memory actually involves two types of belief-dependent processes, synchronic and diachronic. Synchronic belief-dependent processes take beliefs at a

time as inputs and outputs beliefs at that same time.[14] Inference is a synchronic belief-dependent process, as is reconstructive memory. Diachronic belief-dependent processes take beliefs at one time as inputs and output beliefs at some later time. Reproductive memory is a diachronic belief-dependent process.[15]

We must first decide if the process is conditionally reliable to determine whether a belief formed by a belief-dependent process is justified. Secondly, we must determine if the input beliefs are justified. So, for now, supposing that the processes *are* conditionally reliable, our generated memory belief will be prima facie justified just in case the beliefs used to generate it are justified.

Using the distinction between synchronic and diachronic processes above, we can specify *when* input beliefs need to be justified. Input beliefs must be justified at the time they are used as inputs to the process. The input belief in the case of reproductive memory must be justified at the time it is encoded; this is when it is taken as an input to a reproductive memory process. For reconstructive memory, the input beliefs are used at the time the memory is generated, so they must be justified at that time.

The view is as follows:

> S is justified in believing P at time t on the basis[16] of memory if:
>
> 1. S's belief that P was reproduced via a conditionally reliable diachronic memory process and S was justified in believing P at the time of encoding, t_0, or
> 2. S's belief that P was reconstructed via a conditionally reliable synchronic memory process and for all input beliefs, X, used to generate S's belief that P at t, S is justified in believing X at t.

[14]More precisely, the output beliefs are produced shortly afterward given that processes take some time to complete.

[15]Michaelian (2011) suggests process reliabilism as an account of reconstructive memory. His account does not distinguish between synchronic and diachronic belief-dependent processes.

[16]By 'basis,' I mean psychological basis.

I suggested in Section 2 that the justification of reproductive memories is best accounted for by a preservationist account. The reproduced memory is justified only if it was justified at the time it was originally formed. We saw in the previous section that reconstructive memory is generated using traces and our background reproductive memory beliefs. Generated memory is our memory system's best educated guess about the past given the information available. Since reconstructive memory works like a detective, it makes sense that the beliefs it generates are justified in the way that a detective's beliefs would be. Miss Marple's belief that the butler did it is not justified merely because it was the output of Miss Marple's investigative process, no matter how good a detective she is. The evidence and background beliefs she bases her judgment on must themselves be justified.

To take an artificially simple example, let's say that I form an episodic memory that there was a book on the desk in my office last Thursday. Suppose that the memory was constructed based on my beliefs that I typically leave books on my office desk and that last Thursday was a typical day.[17] This belief is formed using a reconstructive memory process, so it is synchronic. My belief that there was a book on my desk last Thursday will be justified only if the beliefs my memory processes use to construct it are themselves justified. Now, suppose that one of these input beliefs, my belief that I typically leave books on my office desk, is purely reproductive. This case falls under clause 1. This belief is justified only if the belief it was based on—that very belief when I initially acquired it—was justified at the time of encoding. Assuming that my memory is generally reliable, my belief about the book will be justified as long as I am justified in believing that I typically leave books on my

[17]Of course, in any actual case, there will be more beliefs used, but as long as they are all justified, the memory belief is justified. Additionally, non-belief factors, such as a memory trace of my experience of the books on the desk might or might not be instrumental in the reconstruction.

desk. I am justified in believing that I typically leave books on my desk only if I was justified in believing that when it was first encoded into memory.

According to this view, reproductive memory preserves justification while any new beliefs constructed by memory are justified inferentially. Note that the content of these constructed beliefs may or may not have been previously entertained or may have been entertained at the time but not stored. This account supplements the preservationist view of reproductive memory with an inferentialist view of reconstructive memory. Unlike the pure preservationist view, this account can accommodate newly constructed memory contents. It does this by making a distinction between semantic and episodic memories based on a modern psychological view of how the processes that produce these beliefs work. In the final section of this chapter. I will defend this view against three principal objections.

3.5 Objections

3.5.1 Reconstructive Memories Aren't Memories

My view supplements preservationism with inferentialism for reconstructive memory beliefs. The preservationist can resist this move by denying that these beliefs really are memories. If they are not memory beliefs, then they do not pose a problem for a theory of the justification of memory beliefs.

It is tempting to view this as simply a terminological dispute. I explained in the introduction that my usage of "memory" isn't ordinary English usage. Perhaps preservationists are merely interested in the class of beliefs that are picked out by the folk concept of memory. They are, of course, free to have such a theory. However, this move is problematic. Since most of our episodic memory beliefs are reconstructed,

denying that these are memory beliefs would mean that many of the beliefs we take to be memories *really* are not. I take that to be an unacceptable cost for a view meaning to adhere to a folk understanding of memory; still, some might be willing to bite this bullet. In that case, reconstructive memory beliefs are not 'memory beliefs', properly speaking. However, regardless of what we want to call them, we still need an account of how these generated 'memory' beliefs are justified.

Many preservationists would be uncomfortable going this route. If by discussing memory beliefs, they mean to be talking about a psychological kind, rather than a folk concept, my argument is even stronger. I have presented psychological evidence for viewing both reconstructive and reproductive memories as memory in the psychological sense. They may be different cognitive systems, but they are both memory systems in the fullest sense; just as vision and hearing are different cognitive systems but are still both perceptual. A theory of the justification of perceptual beliefs must apply to hearing as well as vision.

3.5.2 Skepticism

Avoiding skepticism was one of the desiderata discussed in Section 2. Preservationism avoids the radical skepticism of the inferential view. By combining preservationism with inferentialism, my view also avoids skepticism about reproductive memories. However, it is subject to another kind of skeptical problem.

Huemer (2008) presents a new evil demon style worry about preservationism. Consider a situation where the world was created five minutes ago with an intrinsic duplicate of you. Since she is your duplicate, she has all the same beliefs you do. However, any memory beliefs your duplicate has that were not originally formed in the previous five minutes must be unjustified according to preservationism. The

duplicate has obviously never held the belief before since there was no earlier time for her to have it. However, you are justified in this belief as long as you were justified in forming it. Preservationism would lead to the result that you are justified but your duplicate is not.

Huemer suggests the following "dualistic theory" to solve this problem:

> On this view, a belief is justified full stop if and only if one had an adequate justification for adopting it at some point, and thenceforward one was justified in retaining it. The normal functioning of memory, in the absence of specific reasons for revising a belief, constitutes an epistemically acceptable manner of retaining beliefs.... [C]oming to believe something by seeming to remember it (in the absence of defeaters one is aware of) is an epistemically rational way of acquiring the belief. (1999, 351)

According to Huemer's dualistic view, both you and your duplicate are justified. You are justified because you were justified in adopting it and are justified in retaining it via the normal functioning of memory—the standard preservationist story. Your duplicate is justified because she came to believe it by seeming to remember it—the foundationalist aspect. The dualistic theory allows us to gain the benefits of preservationism while allowing for cases where (apparent) memory generates beliefs.

When it comes to the psychology of reconstructive memory, Huemer's view fails because it faces a version of the problem that his account is meant to avoid. Constructed memories count as coming to believe something by seeming to remember it, and so such memories will be subject to the foundationalist treatment that Huemer endorses. However, constructed memories are sometimes generated by background beliefs that are unjustified. In such cases, Huemer's view faces the same issue that lead him to reject the pure foundationalist theory: we will be justified in accepting whatever we seem to remember constructively, in the same way that the pure foundationalist theory entails blindly accepting whatever we seem to remember reproductively.

To see this, consider a modified 'forgotten carelessness' case. Suppose that you read in *The Onion*, a satirical newspaper, that there is a new trend sweeping graduate student offices. They have started using typewriters to compose their dissertations rather than computers. You realize that it is a joke but believe it anyway. Over time, you retain this belief while forgetting the source. Now, when trying to remember the contents of the office, you use your unjustified stored belief that offices often contain typewriters to construct a belief that there was a typewriter in this office. However, since the constructed memory belief is merely apparent, it is foundationally justified according to Huemer.

3.5.3 Memory is Epistemically Basic

I have been defending a view where memory is a belief-dependent process, and so is not a basic source of justification. Reproductive memories inherit the justification of the original belief while reconstructive memories depend on the justification of the beliefs used to construct them. However, Lyons (2009) provides an account according to which memory is epistemically basic. Just because memory beliefs are psychologically caused by other beliefs does not mean they rely on those beliefs for their justification. This results in foundationalism about memory. Lyons (2009, 144) gives the following account of basic beliefs:

> (B) A belief B is basic for S at t iff B is the output at t of one of S's cognitive systems that
>
> 1. is inferentially opaque,
> 2. has resulted from learning and innate constraints, and
> 3. does not base its outputs on any doxastic inputs at t.

I agree that memory satisfies the first two requirements for basicality. The point of contention is whether it satisfies the third. If outputs of memory are based on doxastic inputs at the time of retrieval, then memory is not basic.

First, let us consider simple reproductive memory belief that is caused by the original belief. It would seem that this memory belief is based on doxastic input.

However, this original belief is no longer accessible. According to Lyons, this disqualifies it for being a basis for the memory belief. What is accessible to the memory system are stored representations that may or may not be beliefs.[18] Here is Lyons' explanation of the problem:

> "The stored representations may be based for Σ [memory system] on the old belief tokens, and the new belief tokens may be based for Σ on these, but it does not follow that the new tokens are based on the old tokens. The basing relation is not transitive. If an output is based for an agent on e[evidence] at t[time], then e is accessible at t to that agent and the same holds for modules. In an ordinary memory case, the belief tokens occurrent at t_e [time of encoding] no longer exist and thus are not accessible to the system at t_r [time of retrieval]; thus they cannot be part of the system's grounds for its outputs. ...So in the standard case of memory, the output beliefs are not based on the input beliefs, even for Σ." (2009, 140)

This basing relation requires that if an output is based on e at t, then e is accessible at t. The original belief is inaccessible at the time of retrieval, so it cannot form the basis for the memory belief. Preservationism requires the denial of this temporal accessibility requirement, and thus Lyons' view rules out preservationism entirely.

Several problems should lead us to abandon Lyons' view. As a purely foundationalist view, it is subject to the same objection discussed in Section 2: it simply gives us too many justified memory beliefs. Interestingly, it also faces the opposite danger, skepticism. Lyons rejects the connection between belief dependence and non-basicality, so memory must be unconditionally reliable for the beliefs it generates to be justified. However, it is reasonable to suppose that memory processes are at great risk for being unconditionally unreliable. For example, we all have many false beliefs that have been stored in memory. This problem becomes more pressing if we are worried about Huemer's example of duplicates in a five-minutes-old-world (discussed in the previous subsection). In that case, almost all of the duplicate's (apparent)

[18]For the time being, let us assume that the stored representations are not beliefs.

memory beliefs are false, and so their memory is unconditionally unreliable if anything is. This means that even their memories about the immediate past (fewer than five minutes ago) are unjustified.

Another problematic feature of this account is that beliefs can become unjustified merely by passing into memory. Take the case of a cult member raised in isolation from outside influence.The cult leaders, whom she trusts implicitly, inculcate many false beliefs in her so that she has more false beliefs than true ones. However, she is at least prima facie justified in these beliefs; she has no way of discovering that the leaders are untrustworthy. Now suppose these beliefs become memories. She has not forgotten anything, and she has not gained any new evidence in the meantime. Since she has more false beliefs than true ones, her memory is unreliable, and these beliefs are now unjustified. This is very counterintuitive.

Now, let us turn to reconstructive memory. Lyons suggests that even though a memory system has access to the stored representations and beliefs used in the reconstruction, the larger subject does not. "All I know about these representations I must either piece together from the outputs of Σ or discover by way of empirical research in the psychology of memory." (2009, 141) Lyons considers beliefs to be defined by a certain functional role, but a representation may play that functional role within a cognitive system without playing that role in the larger organism.[19] Since the constructed memory belief is based on the stored representations, these representations are playing the belief role within the memory system itself. However, Lyons claims, they do not play that role for the agent as a whole. According to Lyons,

[19]We must be "careful to distinguish belief, the global property of an individual, from beliefs, occurrently tokened mental representations with a certain functional role. Whether, in the end, I believe that p might very well depend on how my central systems deal with some mental representation of p, but this is perfectly compatible with the claim that this representation of p is the output of some perceptual module." (2009, 92)

these representations are similar to the complex mathematical assumptions that one's visual system uses to construct visual representations.

While I believe that what Lyons claims is true of some stored representations in a memory system, such as a memory traces, it is not true for all of them. For example, it is not unreasonable to suppose that many of the priors or schemas used to construct episodic memories are things that I believe. For example, the representation that graduate student offices likely contain books is just the kind of thing that I can, and do, believe. In fact, the argument that memory errors are the result of adaptive processes relies on this fact. My office schema is used to construct episodic memories as well as to form expectations of what I might find when entering a new office for the first time. Since these stored representations are used by memory processes and are things that I can explicitly endorse, I would argue that they are beliefs that I, and not merely my memory system, have.

3.6 Conclusion

The purpose of reconstructive memory is not to preserve old beliefs, but to construct new beliefs about the past based on our often limited reproductive memory and our current background beliefs. According to preservationist accounts of memory, a memory belief is justified only when that belief was previously justified. Since reconstructive memory beliefs may not have been held before, or might not have been continuously held, a purely preservationist account of memory cannot accommodate justified reconstructed memory beliefs. Reconstructive memory is an inferential process, and the beliefs it produces are justified in the same way that other inferential beliefs are justified. We can retain a preservationist account of reproductive memory

as long as we add an inferential account of reconstructive memory.

Chapter 4

Can Forgetful Agents be Rational?

4.1 Introduction

According to standard Bayesian conditionalization, forgetting is irrational. In most cases, however, forgetting does not seem like a matter of rationality. Accordingly, many alternatives to conditionalization have been proposed which view forgetting as arational rather than irrational. Nevertheless, the existing literature has mostly overlooked a crucial distinction between two types of forgetting: in addition to forgetting their evidence, an agent might also forget their conditional credences. Is it irrational to forget one's conditional credences or is this also merely arational?

In this chapter, I will canvass several of the proposed alternatives to conditionalization that view forgetting evidence as arational, and explore what verdict they might give for agents who forget their conditional credences. While I agree that forgetting evidence is not irrational, I believe that the jury is still out when it comes to forgetting conditional credences. Thinking carefully about the rational status of forgetting conditional credences can improve our current theories of rationality in the same way forgetting evidence has.

In this chapter, I will start by giving a brief overview of the Bayesian picture and conditionalization. In Section 3, I go over some arguments in favor of conditionalization as a requirement of rationality. Then, in Section 4, I will define two types of

forgetful agents (FAs): those that forget evidence (FA_E) and those that forget conditional credences (FA_C). Section 5 will examine one of the most popular methods of solving the forgetting evidence problem, ur-prior conditionalization. I will explore several options for understanding the nature of ur-priors and how those options affect the rationality of FA_E and FA_C. In Section 6, I will turn to new 'time-slice' versions of ur-prior conditionalization that view rationality as entirely synchronic in nature. I will argue that these are too lax in their treatment of FA_C. The final two sections of the chapter will cover other alternative conceptions of conditionalization. Section 7 focuses on agential views that take conditionalization as a rational commitment that agents adopt by holding conditional credences. Finally, in Section 8 I will shift to process views of conditionalization. I conclude that these views are the most promising in their treatment of FA_C.

4.2 Bayesian Rationality

In this chapter, I will be concerned with the rationality of an agent's credences, or degrees of belief. According to Bayesian epistemology, a rational agent's credences must satisfy certain formal requirements. While actual human agents fall short of these requirements in many ways, it describes idealized rational agents. While Bayesians disagree about the nature and scope of these requirements, the ones discussed here are broadly accepted in their basics.

The basic requirement is that rational credal states can be represented by probabilities. This is called probabilism. More precisely, a credal state must satisfy Kolmogorov's Axioms:

Non-negativity: $Cr(A) \geq 0$, for all A.

Non-negativity: $Cr(A) \geq 0$, for all A.

Normalization: $Cr(\top) = 1$, where $\top$ is a tautology.

Finite Additivity: $Cr(A \vee B) = Cr(A) + Cr(B)$, for all incompatible A, B.

Additionally, constraints are placed on changes in credal states over time. Conditionalization is the most commonly accepted of these.

When E is learned[1] between t_o and t_n, for all H,

Conditionalization: $Cr_{t_n}(H) = Cr_{t_o}(H|E)$

Where $Cr(H|E)$ is the credence in the hypothesis, H, given the evidence, E.

$Cr(H|E)$ is traditionally defined in terms of unconditional credences by the Ratio Formula: $Cr(H|E) = Cr(H \wedge E)/Cr(E)$ when $Cr(E) > 0$ and is undefined otherwise. However, in this chapter, I will take conditional credences to be entities in their own right.[2] This will be important throughout my discussion of forgetful agents. As we will see in the following section, I consider forgetting conditional credences to be a separate issue to forgetting unconditional credences. Perhaps the Ratio Formula, similar to Konglomoverov's Axioms above, describes a condition that must hold between conditional and unconditional credences in order for a credal state to be rational, but I will not be taking one to be defined in terms of the other.

What do we mean when we say that conditionalization is a requirement of rationality? I will consider three possibilities. First, we can view it as a requirement on

[1]E is learned and nothing else. Throughout, I will assume that E is learned with certainty. One might allow that E is learned with less than certainty by moving to Jeffrey Conditionalization. Considering Jeffrey conditionalization would add unnecessary complication, but I think that most of what I say here could be suitably modified.

[2]Hájek (2003) convincingly argues that the Ratio Formula cannot be a definition or even an analysis of conditional credences. In fact, he claims that they ought to be considered primitive.

posterior credal states: they must be related to prior states in a certain way. Alternatively, we can view it as a requirement; not on credal states themselves, but on changes between states. Finally, we can view it as a requirement that an agent takes on by being in a certain credal state.

For all H, E,

State $Cr_n(\cdot)$ is rational given evidence E just in case

$Cr_n(H) = Cr_o(H|E)$, for all H.

Agent By having $Cr_o(H|E) = x$, S is rationally committed to having $Cr_n(H) = x$, for all H if S learns E.

Process The move from $Cr_o(\cdot)$ to $Cr_n(\cdot)$ is rational at t_n just in case $Cr_n(H) = Cr_o(H|E)$, for all H where E is learned at t_n.

According to **State** whether a credal state is rational depends only on the contents of current the credal state and their relation to a prior credal state. This requirement does not put any restrictions on how that relationship comes to be; just that it holds. **Agent** puts a rational requirement on agents by virtue of credal states they held in the past. This requirement is meant to be broad-scoped in the sense of Kolodny (2005). An agent in this situation may satisfy the requirement by either updating their credences accordingly or disavowing their prior conditional credences. Finally, **Process** governs the process of moving from one credal state to another. This views conditionalization as a restriction on the type of process one uses to move from one state to another rather than the mere relation between one state and other as viewed by **State**. Again, **Process** does not itself make any additional claims about the details of that process, as long as the end result is as stated.

Two kinds of forgetful agents (*FA*s) will be the focus of this chapter. FA_E can forget their unconditional credences, while FA_C can forget their conditional credences. *FA*s will be described in more detail in Section 4, but first, let us see how they would be traditionally handled. On the standard, straightforward interpretation of conditionalization, both of kinds of *FA*s will sometimes fail to satisfy conditionalization.

First off, an agent can never forget evidence they have conditionalized on if they only update via conditionalization. This feature of classic conditionalizations is called Independence. After conditionalizing on a some evidence (E), Cr(E) = 1. Since conditionalization cannot reduce an agent's credence in propositions already given credence 1, all later credal states must assign Cr(E) = 1. If the only rational way to change credal states is by conditionalization on learned information, any later state that assigns a different level of credence to E is irrational, or at least arational. FA_E sometimes forgets their past evidence. At that point, their Cr(E) will not be 1. So, FA_E fails to satisfy conditionalization.

Secondly, by conditionalizing, an agent's current credences match their prior conditional credences. If credal states are always and only changed via conditionalization, the agent's conditional credences do not really change. Upon learning E,

$$Cr_{t_n}(H|\cdot) = Cr_{t_o}(H| \cdot \&E).$$

FA_C sometimes forgets conditional credences they held in the past. Because of this, there is no way to guarantee that the equation will hold. Again, the new credal state is not a conditionalization of a prior state and so FA_C irrational.

Sections 5 and 6 will canvass various **State** views. Section 7 presents on **Agent** view and Section 9 will cover **Process** views. It is beyond the scope of this chapter

to evaluate these views overall. Here I will be focused solely on what they have to say about forgetful agents. There is more groundwork to set down first before turning to the individual views. Section 3 gives a brief overview of arguments for adopting conditionalization as a requirement of rationality in the first place. Then, Section 4 will define the two types of forgetful agents in more detail. This section will describe the desiderata according to which the views will be analyzed.

4.3 Why Conditionalize?

Why think that conditionalization is a rational requirement? There are several kinds of arguments that have been given for this. In this section, I will briefly discuss some of the most common arguments.

4.3.1 Diachronic Dutch Book

Diachronic Dutch Book Arguments show that an agent with an update strategy that isn't conditionalization can be subject to a series of bets that result in a sure loss. An agent who updates via conditionalization cannot be exploited in this way. Proponents of this argument make the further claim that being subject to a sure loss is a sign of irrationality; therefore, conditionalization is the only rational update strategy.

This argument originates with David Lewis, but was first published in Teller (1973). We assume that credences reflect an agent's betting behavior so that an agent with d credence in P will pay up to $\$d$ for a bet that pays \$1 if P and nothing otherwise. An agent with a $\mathrm{Cr}(P|Q) = c$ will pay up to $\$c$ for such a bet. We also assume that the bookmaker knows the agent's current credences and their update strategy, but have no special knowledge of the situation that the agent doesn't have.

Suppose that our agent, S, for some A, B, plans to assign credence y to A after receiving evidence B where $y \neq$ prior $\mathrm{Cr}(A|B)$. S will accept a series of bets where they will receive a sure loss.

I won't go into the details of the bets here. What's important for my purposes is that the Dutch book can only be made if S has an update strategy precise enough that the bookie knows in advance whether their updated credence will be higher or lower than their prior conditional credence. Additionally, S actually has to carry out the strategy.

Reflection

The *special reflection principle* Van Fraassen (1995) is an additional synchronic constraint intended as a more general alternative to the diachronic constraint of conditionalization. We can then consider an alternative form of Bayesianism that doesn't impose any constraints (like conditionalization) on changing beliefs, but only constraints on what beliefs may be held at a given time.

$$Cr_0(A|Cr_t(A) = x) = x \text{ when defined}$$

Alternately, Briggs (2009) proposes *Qualified Reflection*:

$$Cr_0(A|Cr_t(A) = x) = x \text{ as long as for all B,}$$

1. $Cr_0(Cr_0(A|E) = Cr_1(A|E)) = 1$ and
2. $Cr_(0)(E|Cr_1(B) = 1) = 1$

Qualified Reflection follows from the Kolmogorov Axioms stated above as long as

the following assumptions are met:

1. $Cr_0(Cr_0(A|B) = r) = 1$ iff $Cr_0(A|B) = r$

2. The possible total evidence (B) between t_0 and $t_(1)$ form a partition $\{B_1, B_2, \ldots, B_n\}$.

3. "All agents can be reasonably certain that conditionalization is the right updating procedure."

> "As a logic, Bayesian decision theory ought to be useful for inference and planning. But, some sort of intrapersonal coherence is necessary for inference and planning; an agent who conducts his or her epistemic life correctly will have earlier and later selves that cohere better than a pair of strangers. The sort of diachronic coherence in question should not be so strong as to demand that agents never change their beliefs. But it should be strong enough to bar agents from adopting belief revision policies which lead to changes that are senseless or insupportable by their current lights."(67)

Again, the focus is on the belief revision conception of conditionalization.

Two problems Dutch books might show: incoherence and self-doubt. "Agents are self-doubting when they suspect themselves of having incoherent beliefs." (78–79)

"A set of bets reveals incoherence just in case at every possible world, the buyer of those bets loses more than he or she wins. But a set of bets counts as a Dutch book just in case at every possible world where the agent's beliefs condone the bets, the buyer of those bets loses more than he or she wins. So every set of bets that reveals incoherence counts as a Dutch book, but not every Dutch book reveals incoherence." (80)

4.3.2 Accuracy

Greaves and Wallace (2006) argue that conditionalization is the only rule that maximizes expected epistemic utility. They make it very clear that they are discussing belief revision.

> "The decision problems with which we will be concerned take the following form. The agent begins in some fixed belief state (that is, he holds some fixed initial credence distribution). He knows that he is about to receive some new piece of information, from among a fixed range of possibilities. Before receiving the information, he chooses an updating policy: that is, he specifies, for each of the possible pieces of new information, how he will change his credence distribution if that turns out to be the information that he does in fact receive. The decision he has to make is the choice of an updating policy. " (Greaves and Wallace (2006) 610)

They assume that a rational agent will choose the available act that maximizes their expected epistemic utility. The acts available to an agent are functions from pieces of information (evidence) to credence distributions ($a(s) = p$). This is the probability distribution that the agent would adopt upon relieving that information. Epistemic utility[3] is a function from states of the world and credence distributions to a numerical value. The utility of credence distribution p in state s is $U(s, p)$. The expected epistemic utility of an act is the sum of the utilities in each possible state weighted by the agent's prior credence distribution over those states (p*).

$$\textbf{Expected Epistemic Utility: } EU(a) = \sum_{s \in \mathcal{S}} p^*(s) \times U(s, a(s))$$

Greaves and Wallace prove that conditionalization is the only action that maximizes the expected epistemic utility given certain kinds of epistemic utility functions.

[3]The utility function is a scoring rule. See Joyce (2009).

Since rational agents choose the action that maximizes their expected utility, conditionalization is the only rational action. I do not dispute their proof. I want to draw attention to two aspects of their set up of the decision problem. First, as mentioned above, they are focusing on the choice of an update policy. Second, the expected utility is based on the agent's prior probability distribution. This makes sense because the prior probability distribution is the one the agent has when contemplating which update policy to adopt.

4.4 Forgetful Agents

In the Bayesian framework from the previous section, there are two main components that make up your credal state; conditional credences and unconditional credences. When an agent updates via conditionalization, the agent's prior credences affect their new credences based on new evidence received. It makes sense then that there are two kinds of forgetfulness that might affect whether a forgetful agent can satisfy conditionalization. An agent might forget prior unconditional credences; most interestingly, evidence that they have conditionalized on in the past. I will call this hypothetical agent FA_E. Secondly, an agent might forget their previously held conditional credences. I will call this agent FA_C. For now, unless stated otherwise, I will assume that *FA*s credences at any given time satisfy the Kolmogorov Axioms as above and participate in instantaneous, costless deliberations. *FA*s are therefore still considerably idealized. We humans are not forgetful agents in this sense, but considering them can help us isolate the effects of our imperfect memory from the effects of our many other epistemic shortcomings.

As mentioned in Section 2, conditional credences have standardly been thought of

as being defined in terms of unconditional credences. However, I wish to keep open to the possibility that they are individual entities in their own right. Throughout the rest of this chapter, I will be assuming that conditional credences and unconditional credences are each primitive contentful states; perhaps with something like the ratio formula governing rational connections between them.[4] Under this assumption, we can and should consider the rational status of FA_E and FA_C separately. This is what I aim to do in this chapter.

Throughout the rest of this chapter, we will see that it is commonly agreed that FA_E can be rational. As Williamson (1998) famously puts it, "[f]orgetting is not irrational, it is just unfortunate." (98) By 'forgetting' here he is talking about forgetting past evidence. I have no desire to break with the status quo in this respect. There has been much less written about FA_C. I will not be defending a settled position on FA_C here. Rather, I think there may be several reasonable positions to hold in this case. My aim in this chapter is primarily to explore in detail what existing views of conditionalization have to tell us about the rationality of forgetting conditional credences. I will give mild preference to views that leave open whether or not FA_C could be rational, as, in the end, I think it may come down to details about what conditional credences are like for that agent. For example, a FA_C with mental states somewhat like ours may not be rational, while other kinds of agents might be.[5]

For the remainder of this chapter, I will discuss alternative views of conditionalization and how they treat both kinds of *FA*s. Even on the **State** interpretation,

[4]Here I hold a view similar to that in Titelbaum (2013) which takes "conditional credence to be a genuine mental state (an attitude towards an ordered pair of propositions) capable of being elicited in various ways, such as by asking an agent her confidence in a proposition given a supposition."(54)

[5]For example, in Chapter 2 of this dissertation I argue that memories of conditional credences tend to differ greatly from unconditional credences in their nature and justificatory status. This might make a difference in what is rational for creatures like us in this respect.

there have been suggestions for understanding prior conditional credences as well as unconditional credences in a synchronic rather than diachronic way. Ur-priors make conditional credences a synchronic matter. Likewise, total evidence makes unconditional credences synchronic rather than diachronic. The versions of ur-prior and total evidence views can then be combined to form more or less synchronic views of conditionalization. While they all agree that FA_E can be rational; they give differing judgments of FA_C. After evaluating these views, I will turn my attention to the other interpretations of conditionalization that I am calling **Agent** and **Process**.

In the following section, I will explore different ways of understanding conditional credences and how they relate to conditionalization. While these views do not explicitly consider the fate of FA_C, I will attempt to draw out the implications as much as possible. Then I will evaluate these views on whether their treatment of FA_C seems reasonable.

4.5 Ur-Priors

If you recall from Section 2, the **State** interpretation of conditionalization says:

$Crn(\cdot)$ is rational given evidence E just in case

$$Cr_n(H) = Cr_o(H|E), \text{ for all } H$$

One question that arises based on this definition is just what these prior conditional credences, $Cr_o(H|E)$, are supposed to be. The standard interpretation of conditionalization takes them to be credences actually held by the agent in question at some time (immediately) prior to the update. However, recently it has become

popular to understand prior conditional credences as ur-priors. Rather than considering updates happening sequentially, we take the ur-prior conditional credences and apply them to the agent's current total evidence.[6]

Ur-Prior Conditionalization: $Cr_n(\cdot)$ is rational given the agent's current total evidence, E, just in case

$$Cr_n(H) = Cr_{up}(H|E), \text{ for all } H.$$

Where *up* are the agent's ur-priors.

This section will focus on three different kinds of things one might have in mind with ur-priors; personal ur-priors represent an agent's personal conditional credences at a time, functional ur-priors are an abstract functions that allow an agent's updates to satisfy conditionalization, and evidential standards ur-priors represent either the evidential standards of an agent or some objectively correct evidential standards. The view of ur-priors adopted will affect the rational standing of Forgetful Agents under ur-prior conditionalization.[7]

4.5.1 Personal

As a first suggestion, one might take these ur-priors to be the agent's actual original conditional credences. Typically the agent's credences before gaining any evidence whatsoever. If there is no such time, then ur-priors are the credences of the agent at the earliest available time. Of course, there is so far no guarantee that an agent's

[6]Most of the accounts of Ur-prior conditionalization have agents conditionalize on their total evidence at a time. However, I want to consider these two aspects separately as much as possible. Total evidence will be explored in Section 6.

[7]Throughout this section, I will use classifications adapted from those in Meacham (2016) with some simplifications.

original credences are themselves rational. To get a normative account of conditionalization on this view, one must find the coherent ur-priors that are the closest fit to the agent's actual priors.[8]

Setting aside any problems this type of view might face, let us consider how it treats forgetful agents. On a personal conception of ur-priors, the agent's conditional credences are determined in the absence of evidence (as much as this is possible). At any given time, these priors are then used to update on the agent's total evidence at that time.

First, let us look at FA_E who forgets evidence upon which they have previously updated. Recall that they get into trouble with traditional conditionalization because once they update on a piece of evidence, their credence in that evidence must remain 1 forever. Ur-conditionalization on personal priors requires them to update on their total evidence based on their 'original' conditional credences instead. We will discuss the idea of total evidence in more detail in Section 6. For now, the important thing is that personal ur-priors and total evidence combined allow FA_E to satisfy Ur-conditionalization.

FA_C will clearly not satisfy conditionalization in this case. When FA_C forgets their conditional credences, their ur-priors have been forgotten. FA_C cannot be expected to update using priors they no longer have access to (except perhaps by using the original credences coincidently). I take this kind of view to be overly harsh on FA_C. Their ur-priors are merely conditional credences held at the earliest possible time. There is no compelling reason given for why rationality requires an agent to stick to their original conditional credences for all time.

[8]Since I am only considering agents that satisfy the synchronic norms given in Section 2, we can ignore this issue here.

4.5.2 Functional

Alternatively, ur-priors might be understood functionally. That is, the ur-priors are just whatever plays the functional role of priors. "On this understanding, the norm just requires that there exist some function–call it the ur-prior function—that plays a certain role with respect to a subject's beliefs. In particular, the norm requires that there exist some function, *up*, such that her credences at every time are equal to up conditional on her evidence at that time." (Mecham 471) In other words, to discover an agent's functional ur-priors you look at all the times the agent updated their beliefs, and find the set of priors that would make those updates satisfy conditionalization. Depending on how the agent updates their credences, there may be several such functions, or there may be none. As with personal ur-priors, there may be additional rational norms that further constrain appropriate functions.

How does FA_C fare here? This is a much more difficult question to answer. In order to (non-vacuously) satisfy ur-prior conditionalization, there must be at least one function that can describe all of their updates. The functional understanding does not require that an agent is at all aware of their ur-priors, so the fact that FA_C forgets conditional credences does not automatically rule this out. They are free to use any kind of updating procedures they like, and it just might work out in such a way that we could draw up an appropriate ur-prior function. Consider the case where FA_C always updates based on what they currently take their conditional credences to be, and those priors change over time. In this situation, there will not be an ur-prior function that captures all of those updates. In this case, they will fail to satisfy ur-prior conditionalization.

In order to solve for an ur-prior function, the agent's unconditional credences are taken as given. This makes it difficult to determine the rationality of FA_E as well.

As it stands, an agent might be considered irrational if their unconditional credences make it impossible to fit a coherent ur-prior function to them.

4.5.3 Evidential Standards

Perhaps the most popular conception of ur-priors that they represent evidential standards. Meacham (2016) provides several suggestions for evidential standards. Evidential standards might be whatever an agent considers reasonable independently of their evidence. Alternatively, they might be positions the agent is somehow committed to, leaving the nature of the commitment involved open to many different interpretations on its own. Evidential standards might merely represent the agent's dispositions to form beliefs. Finally, I am going to consider a view of ur-priors as some objectively correct evidential standards as falling under this category as well.

One of the most famous examples of an evidential standards view is found in Williamson (1998).

"Let P be the prior probability distribution, e_w the conjunction of all old and new evidence for some individual or community S in a circumstance w, and $P_w(p)$ the evidential probability of a proposition p for S in w. The proposal is that P_w is the conditionalization of P on e_w,

$$\textbf{ECOND} P_w(p) = P(p|e_w) = P(p\&e_w)/P(e_w)(P(e_w) > 0)$$

The implications for FA_E are entirely straightforward. Williamson's version of conditionalization always applies to an agent's current total evidence in a situation. Unconditional credences held in the past play no role whatsoever. There is nothing in the way of FA_E being rational.

The prior probability distribution here is a measure of "something like the intrinsic plausibility of hypotheses prior to investigation." (211) It is not meant to be an agent's actual credal state at any time. According to Williamson, there is only one appropriate ur-prior. This means that for any given total evidence, there is only one rational credal state. This is called the Uniqueness Thesis.

Since this hypothetical ur-prior is something an agent might never actually hold (all that matters is that they update as if it was their prior credal state, similar to the case with functional ur-priors), the result is a mixed bag for FA_C. There are three possibilities here. First, if their conditional credences were not the appropriate ones, and forgetting them leads them to update according to ECOND, then forgetting actually makes them rational when they were not before. Second, if they happened to have the appropriate priors to start and subsequently forget them, updating on their current priors will be irrational. Finally, if forgetting merely moves them from one inappropriate set of priors to another, they remain irrational throughout.

Finally, let us turn to evidential standards that are subjective to the agent. There are two options for this type of view. First, if the ur-priors are merely meant to reflect standards that the agent is in some sense committed to, the agent need not even be aware of them. All that matters is that the agent's belief changes conform to them. Secondly, they might be taken to be actual mental states of the agent in question.

Taking the first option gives similar results as in the objective standards case above. The focus on total evidence puts FA_E in the clear. FA_C's fate is less clear. It will depend on whether forgetting results in conditional credences that align with the appropriate evidential standards.

I find that taking ur-priors to be mental states representing the agent's evidential standards to be the more interesting option. In the following section, I will describe

some of these views and discuss how they treat FA_E and FA_C.

4.6 Time-Slice Rationality

The standard interpretation of conditionalization governs updates that happen cumulatively over time. An agent receives a piece of evidence and conditionalizes on that. This generates new conditional credences which will be used the next time an agent gets a new piece of evidence. This results in the monotonicity discussed in Section 3. Once a piece of evidence is conditionalized on, it can never have a credence less than 1; this leads to an obvious problem for Forgetful Agents.

Instead, as we saw in the previous section, ur-prior conditionalization updates on all of an agent's evidence at once using ur-priors.

Ur-Prior Conditionalization: $Cr_n(\cdot)$ is rational given the agent's current total evidence, E, just in case

$$Cr_n(H) = Cr_{up}(H|E), \text{ for all } H.$$

Where *up* are the agent's ur-priors.

Defenders of Ur-prior conditionalization differ on what an agent's total evidence is. Since the nature of evidence is entirely too large of a topic to cover in this chapter, I will be only presenting these views without discussing any of the arguments in favor of one over the other. The focus here will be solely on how they handle Forgetful Agents.

There has recently been a push towards a time-slice epistemology. These views are partly motivated by problems with traditional diachronic norms like conditionalization. They also mostly tend to be inspired by internalism to one degree or another.

These views aim to allow forgetful agents (in particular FA_Es) to be completely rational.

Moss (2013) explains her time-slice theory of rationality as follows:

> At a first pass, we define this theory as the combination of two claims. The first claim: what is rationally permissible or obligatory for you at some time is entirely determined by what mental states you are in at that time. This supervenience claim governs facts about the rationality of your actions, as well as the rationality of your full beliefs and your degreed belief states. The second claim: the fundamental facts about rationality are exhausted by these temporally local facts." (1)

Similarly, Hedden (2015) provides an internalist motivation for his version of time-slice rationality. ”Internalists hold that facts about the external world or about the reliability of your perception do not affect what you ought to believe, except insofar as they affect your mental state. And by the same token, they should hold that facts about how you were in the past or about the reliability of your memory do not affect what you ought to believe, except insofar as they affect your present mental state. This means that we should accept Synchronicity.” (5) Synchronicity is the claim that “what attitudes you ought to have at a time does not depend on what attitudes you have at other times.”(4)

Time-slice rationality defenders differ on what counts as a mental state. For example, Silins (2005) defends the view that “[n]ecessarily, if A and B are internal twins, then A and B have the same evidence.” (376) Where A an B count as internal twins iff they have all the same non-factive mental states. While Williamson (1998), claims that evidence is the agent’s knowledge (E=K). Where knowledge is a factive mental state. Whatever the details, these views share the verdict that an agent’s total evidence supervenes on their current mental states. FA_E) is entirely vindicated on this view of total evidence.

Typically, time-slicers take a mental-state view of total evidence to be enough to

establish ur-prior conditionalization as an entirely synchronic norm. However, there is another important component to consider; the ur-priors themselves. To determine the fate of time-slice epistemology and FA_C, we must return to the theories of ur-priors from the previous section.

As mentioned in the previous section, we might take ur-priors to be current mental states of an agent that represent their evidential standards. This gives us an entirely synchronic version of conditionalization. The agent's evidence, as well as her evidential standards, all supervene on her current mental state. The agent's past mental states and any other external factions play no direct role.[9] Nothing about an agent's past credal states factor into the rationality of their current credences. Rationality is entirely a synchronic matter on these views, so it is no surprise that Forgetful Agents can still be considered rational. It might be better in some sense if we did not forget information, but whatever sense that is has nothing to do with rationality.

Time-slice views seem intuitive when we are only considering FA_E. This is unsurprising as allowing for forgetting evidence is a focal case for such views. However, one might worry that this view is altogether too lenient on FA_C. To see this, let us return to the issues that arise when we deny the Uniqueness Thesis from Section 5.

If the Uniqueness Thesis fails, there are multiple acceptable ur-priors in at least some situations. Such a view is subject to the question of how to treat shifts between different sets of acceptable ur-priors over time. Once an agent settles on one of the acceptable ur-priors, must they stick with it? Or are they free to move between them at will? Time-slicers are committed to having no restrictions on changes in ur-priors over time; only the current conditional credences matter for rationality.

For example, Moss (2015) explains that

[9]They may play an indirect role by influencing the agent's current mental state, or due to externalism about mental content, but that is the limit of their influence.

> "[i]n a decision situation, an agent must act to maximize expected value according to the precise mental state she identifies with. But there is no rule of rationality saying that an agent cannot change which mental state she identifies with. If an agent does start to identify with another precise mental state, she must then act to maximize expected value according to that precise mental state. This is one way in which a rational agent might 'change her mind,' colloquially speaking. She may start to identify with another precise mental state even if she has already made decisions that did not maximize expected utility according to that state." (673)

This quote makes it clear that FA_C is entirely off the hook. An agent need not be held to their previous conditional credences if they have been forgotten or even if the agent merely changes their mind. This is true even if the changes result in failures to maximize expected utility.

4.7 Agential Views

Recall the **Agent** conception of conditionalization from Section 2:

Agent: By having $Cr_o(H|E) = x$, S is rationally committed to having $Cr_n(H) = x$, for all H if S learns E.

This is a way of understanding conditionalization as a commitment agents take on by holding certain conditional credences. Now, there is a distinction between what we plan to do and what we actually do. Many things can get in the way of executing even the best-laid plans. In non-epistemic decisions, an agent might receive some new information, their desires might change, they might suffer from weakness of will, or they might forget what the plan was. In most cases, we assume that these things have not happened. Once we focus on Forgetful Agents, this is a serious problem. If an agent forgets their past conditional credences, they do not remember what new credences they planned to have. If an agent forgets their commitments, it seems they have little choice but to fail to satisfy them.

Suppose FA_C plans to conditionalize on some past conditional credences but subsequently forgets them. Instead, they update on their current conditional credences. Is there anything irrational about that? The question of what makes conditional credences rational in the first place arises yet again. Why is it irrational to change them once I have picked? If there is just one rational prior credal distribution, then forgetting is irrational (assuming we are at the right one to start with). If there is some set of rational priors, what's wrong with going from one to another? If any (coherent) prior is rational, again why not move between them? A book cannot be made against such an agent. The bookie will not know exactly what they will do. It will not maximize expected epistemic utility by the agent's past lights, but it will by their lights at that later time. This mirrors the view of the time-slicers in the previous section.

Again, this is the situation: Currently $cr(H|E) = x$, but there are other credences the agent could (rationally) have, say $cr(H|E) = y$. Once an agent 'picks' $cr(H|E) = x$, are they forced to stick to it? Does that make it irrational for them to later have $cr(H|E) = y$ and update accordingly? Even if they forget what their earlier credence was?

Bratman (1987) gives an iconic case where it seems irrational to change a practical plan after it has been chosen. There are two routes I could take to San Francisco; 101 or 280. To get to Route 101, I need to make a right at a particular intersection and to get to 280, I need to go left. I consider both to be equally good but arbitrarily form the intention to take Route 101. Now, I am at the intersection. According to Bratman, it's irrational for me to now make a left even though it would not have been irrational had I not formed that intention. (23)

If a view like Bratman's is correct, then there is something irrational about failing

to stick to a plan after it has already been made. We can adapt this to conditional credences as well if they are plans. The idea would be that by having certain conditional credences, the agent is planning to update in a certain way in the future. By making this plan, the agent is taking on an additional rational requirement. This would run counter to the time-slice view of the previous section. Planning is one way that past mental states (past conditional credences in this case) can affect the rationality of an agent's current credal state. However, according to **Agent**, unconditional credences do not carry the same kind of commitment.

On this view then, FA_E is rational, but FA_C might not be. This view is not as harsh on FA_C as some of the others considered so far. It leaves open the possibility that there are certain situations where it would be rational to change your plans. In those situations, FA_C may still be rational.

4.8 Process Views

Finally, let us examine the process interpreation of conditionalization:

Process: The move from $Cr_o(\cdot)$ to $Cr_n(\cdot)$ is rational at t_n just in case $Cr_n(H) = Cr_o(H|E)$, for all H where E is learned at t_n.

Recently, Titelbaum (2015) and Podgorski (2016) have used process views to defend diachronic norms against the synchronic norms from Section 6. According to Podgorski (2016), "[d]iachronic norms govern *processes*, temporally extended, causally continuous patterns of mental states" (5) Similarly, Titelbaum (2015) argues that "Reasoning is a crucial rational activity; being causal, it extends over time; instability of belief would vitiate reasoning's efficacy."(682) Diachronic norms govern processes like belief change or reasoning. This is a thicker concept than just relations between

states. Since remembering (and forgetting) is also a causal, temporally extended process, this type of view seems like a promising way to understand FA_C.

Titelbaum (2013, 154) provides an interesting case:

> "Baseball: The A's are playing the Giants tonight, and in the course of a broader conversation A's announcers Ray and Ken turn to the question of who will win the game. They agree that it's a tough matchup to call: the Giants have better pitching, but the A's have a more potent offense; the A's have won most of the matchups in the past, but the A's are weaker this year than usual. All in all, it seems like a reasonable person could go either way. Nevertheless, Ken asks Ray what he thinks, and Ray says 'I'm not certain either way, but I think it'll be the A's.' Ray then goes on to discuss how an A's win might affect the American League pennant race, etc. ...
> Five minutes later, Bill comes in and asks Ray who he thinks will win tonight's game. Ray says, 'I'm not certain either way, but I think it'll be the Giants.' "

Titelbaum argues that Ray's behavior is at least 'rationally problematic.' The case itself does not specify a reason for Ray's change in opinion. However, for our purposes here, let us suppose that there are two sets of conditional credences available to Ray, both of which are equally rational. When giving his opinion to Ken, Ray updates on one set of conditional credences and his total evidence at that time. Then he promptly forgets about his chosen conditional credences. When asked by Bill, Ray uses the other set. He updates on the same evidence as before but forms the opposite opinion. In this version of the story, Ray is a FA_C.

On Titelbaum's view, a FA_C like Ray is not necessarily irrational in the strictest sense of the word, but there is still something not quite right about them. "At best, we get that it's rationally harmful when an agent's attitudes shift around in certain kinds of cases. The rational verdict is evaluative, Not Prescriptive" (685). Forgetting is intuitively different from consciously changing your mind. When you realize you have changed your mind, you will tend to go back and start any deliberations over using the new beliefs. If you forget or unconsciously change your mind, you cannot do that. This may result in bad reasoning if you were using the changing beliefs at

the time. This type of process view allows for varying verdicts for FA_C, depending on the details of any particular case.

4.9 Conclusion

Can forgetful agents be rational? In this chapter I have argued that there are two questions here: 'Can it be rational to forget prior evidence?' and 'Can it be rational to forget prior conditional credences?' Conditional credences and evidence should be evaluated as separate entities, though this distinction is often neglected. Throughout this chapter, we have examined different views of Bayesian rationality and found that they offer a wide variety of answers to the second question. Meanwhile, they all agree that the standard view of conditionalization is overly harsh on agents who forget their evidence. However, I have suggested here that while these views may get the correct answer for agents who forget their evidence, most are either too harsh or too forgiving of agents that forget their conditional credences. I believe that the most reasonable solution lies in some version of a process view. Regardless of one's preferred account of rationality, I hope that the possibility of forgetting conditional credences will receive more of the attention it deserves in the future.

Bibliography

Alter, A. L. and Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and social psychology review*.

Audi, R. (1995). Memorial justification. *Philosophical Topics*, 23(1):31–45.

Bernecker, S. (2009). *Memory: A philosophical study*. Oxford University Press.

Bernecker, S. (2010). *Memory: A philosophical study*. Oxford University Press.

Binder, J. R. and Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in cognitive sciences*, 15(11):527–536.

Bratman, M. (1987). *Intention, plans, and practical reason*. Harvard University Press, Cambridge, Mass.

Brewer, W. and Treyens, J. (1981). Role of schemata in memory for places. *Cognitive Psychology*, 13(2):207–230.

Briggs, R. (2009). Distorted reflection. *Philosophical Review*, 118(1):59–85.

Cabeza, R. and Moscovitch, M. (2013). Memory systems, processing modes, and components: functional neuroimaging evidence. *Perspectives on Psychological Science*, 8(1):49–55.

Cheng, S. and Werning, M. (2016). What is episodic memory if it is a natural kind? *Synthese*, 193(5):1345–1385.

Cleary, A. M. (2008). Recognition memory, familiarity, and déjà vu experiences. *Current Directions in Psychological Science*, 17(5):353–357.

Cubelli, R. (2010). A new taxonomy of memory and forgetting. In Salla, S. D., editor, *Forgetting*. Psycholoy Press.

De Brigard, F. (2014). Is memory for remembering? recollection as a form of episodic hypothetical thinking. *Synthese*, 191(2):155–185.

De Brigard, F. and Giovanello, K. (2012). Influence of outcome valence in the subjective experience of episodic past, future, and counterfactual thinking. *Consciousness and Cognition*, 21(3):1085–1096.

Debus, D. (2010). Accounting for epistemic relevance: A new problem for the causal theory of memory. *American philosophical quarterly*, 47(1):17–29.

Gabrieli, J. D. (1998). Cognitive neuroscience of human memory. *Annual review of psychology*, 49(1):87–115.

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the drm illusion. *Memory & Cognition*, 38(7):833–848.

Goldman, A. I. (1979/2008). What is justified belief? In Sosa, E., Kim, J., Fantl, J., and McGrath, M., editors, *Epistemology: An Anthology*, chapter 29, pages 333–347. Wiley, 2nd edition.

Goldman, A. I. (1999). Internalism exposed. *The Journal of Philosophy*, pages 271–293.

Greaves, H. and Wallace, D. (2006). Justifying conditionalization: Conditionalization maximizes expected epistemic utility. *Mind*, 115(459):607–632.

Greenberg, D. L. and Verfaellie, M. (2010). Interdependence of episodic and semantic memory: Evidence from neuropsychology. *Journal of the International Neuropsychological Society*, 16(5):748–753.

Hájek, A. (2003). What conditional probability could not be. *Synthese*, 137(3):273–323.

Hassabis, D. and Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in cognitive sciences*, 11(7):299–306.

Hedden, B. (2015). Time-slice rationality. *Mind*, 124(494):449–491.

Hemmer, P. and Steyvers, M. (2009). Integrating episodic and semantic information in memory for natural scenes. In *Proceedings of the 31th annual conference of the cognitive science society*, pages 1557–1562. Cognitive Science Society Austin, TX.

Huemer, M. (2008). The problem of memory knowledge. In Sosa, E., Kim, J., Fantl, J., and McGrath, M., editors, *Epistemology: An Anthology*, chapter 58, pages 868–875. Wiley, 2nd edition.

Intraub, H. and Richardson, M. (1989). Wide-angle memories of close-up scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(2):179.

Joyce, J. M. (2009). Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In *Degrees of belief*, pages 263–297. Springer.

Jung, C. (1968). *Man and his symbols*. Dell Pub Co.

Klein, S. B. (2014). What memory is. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(1):1–38.

Klein, S. B. and Nichols, S. (2012). Memory and the sense of personal identity. *Mind*, pages f677–702.

Kolodny, N. (2005). Why be rational? *Mind*, 114(455):509–563.

Koriat, A. (2007). Remembering: Metacognitive monitoring and control processes. In *Science of Memory: Concepts*. Oxford University Press.

Lackey, J. (2005). Memory as a generative epistemic source. *Philosophy and phenomenological research*, 70(3):636–658.

Loftus, E. F. (1975). Leading questions and the eyewittness report. *Cognitive Psychology*, 7:550–572.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, 12(4):361–366.

Lyons, J. (2009). *Perception and Basic Beliefs: Zombies, Modules, and the Problem of the External World*. Oxford University Press.

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

Mazzoni, G. and Memon, A. (2003). Imagination can create false autobiographical memories. *Psychological Science*, 14:186–188.

Meacham, C. J. (2016). Ur-priors, conditionalization, and ur-prior conditionalization. *Ergo, an Open Access Journal of Philosophy*, 3.

Michaelian, K. (2011). Generative memory. *Philosophical Psychology*.

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

Moss, S. (2013). Time-slice epistemology and action under indeterminacy. *Oxford Studies in Epistemology*, 5.

Moss, S. (2015). Credal dilemmas. *Noûs*, 49(4):665–683.

O'Connor, A. R., Lever, C., and Moulin, C. J. A. (2010). Novel insights into false recollection: A model of déjà vécu. *Cognitive Neuropsychiatry*, 15(1-3):118–144.

Plantinga, A. (1993). *Warrant and Proper Functioning*. Oxford University Press.

Podgorski, A. (2016). A reply to the synchronist. *Mind*, 125(499):859–871.

Pollock, J. (1974). *Knowledge and Justification*. Princeton University Press.

Robins, S. K. (2016). Misremembering. *Philosophical Psychology*, 29(3):432–447.

Roediger, H. and McDermott, K. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology-learning memory and cognition*, 21(4):803–814.

Salvaggio, M. (2018). The justification of reconstructive and reproductive memory beliefs. *Philosophical Studies*, 175(3):649–663.

Schacter, D. L. (2007). Memory: Delineating the core. In *Science of Memory: Concepts*. Oxford University Press.

Schacter, D. L., Guerin, S. A., and Jacques, P. L. S. (2011). Memory distortion: an apaptive perpective. *Trends in Cognitive Sciences*, 15(10):467–474.

Schacter, D. L. and Tulving, E. (1994). What are the memory systems of 1994? In Schater, D. L. and Tulving, E., editors, *Memory Systems 1994*, chapter 1, pages 1–38. MIT Press.

Schroer, R. (2008). memory foundationalism and the problem of unforgotten carelessness. *Pacific Philosophical Quarterly*, 89(1):74–85.

Senor, T. D. (1993). Internalistic foundationalism and the justification of memory belief. *Synthese*, 94(3):453–476.

Silins, N. (2005). Deception and evidence. *Philosophical Perspectives*, 19(1):375–404.

Skinner, B. (1983). Intellectual self-management in old age. *American Psychologist*, 38(3):239–244.

Spanò, G., Intraub, H., and Edgin, J. O. (2017). Testing the Boundaries of boundary extension: Anticipatory scene representation across development and disorder. *Hippocampus*.

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

Sutton, J. and Windhorst, C. (2009). Extended and constructive remembering: Two notes on martin and deutscher. *Crossroads*, 4:79–91.

Teller, P. (1973). Conditionalization and observation. *Synthese*, 26(2):218–258.

Titelbaum, M. G. (2013). *Quitting certainties: a Bayesian framework modeling degrees of belief*. Oxford University Press.

Titelbaum, M. G. (2015). Continuing on. *Canadian Journal of Philosophy*, 45(5-6):670–691.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, 53:1–25.

Tulving, E. (2005). Concepts of memory. In Tulving, E. and Craik, F., editors, *The Oxford Handbook of Memory*, chapter 2, pages 33–3. Oxford University Press.

Tulving, E. (2007). Coding and representation: Searching for a home in the brain. In *Science of Memory: Concepts*. Oxford University Press.

Van Fraassen, B. C. (1995). Belief and the problem of ulysses and the sirens. *Philosophical studies*, 77(1):7–37.

Williamson, T. (1998). Conditionalizing on knowledge. *The British journal for the philosophy of science*, 49(1):89–121.

Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press.