The Dissertation Committee for Steven Patrick James Certifies that this is the approved version of the following dissertation:

A Generative Epistemic Theory of Remembering

Committee:

Michael Tye, Supervisor

Sven Bernecker

Ray Buchanan

Josh Dever

Adam Pautz

Mark Sainsbury

David Sosa

A Generative Epistemic Theory of Remembering

by

Steven Patrick James, B.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2014

To the memory of my parents, Francis and Marilyn.

Acknowledgements

I owe thanks to many people. Portions of this project were presented at a meeting of the Central APA, as well as conferences at USC, Texas Tech, Virginia Tech, University of Western Ontario, University of Waterloo, University of Glasgow, and colloquia and dissertation seminars at the University of Texas at Austin. I thank the audiences as these events for their helpful feedback. A portion of Chapter 4 appears in very similar form in print at *Synthese* under the title "Hallucinating Real Things" (DOI) 10.10007/s11229-014-492-4, and I thank them for permission to include the work in this dissertation. Additionally, I am deeply indebted to all of the following individuals with whom I have had many helpful conversations: Sven Bernecker, Brian Cutter, Sinan Dogramaci, Jeremy Evans, Alex Grzankowski, Cory Juhl, Hans Kamp, Al Martinich, Michelle Montague, Jo Painz, Adam Pautz, Bryan Pickel, Kate Ritchie, Miriam Schoenfield, Jason Schukraft, David Sosa, and Galen Strawson. I am especially indebted to Ray Buchanan, Josh Dever, Mark Sainsbury, and Michael Tye, who have provided extremely valuable feedback at every stage of this project.

An Epistemic Theory of Remembering

Steven Patrick James, Ph.D.

The University of Texas at Austin, 2014

Supervisor: Michael Tye

This dissertation is about the nature and epistemic significance of remembering. Recent philosophical work has exploited the constructive nature of memory to weaken its relationship to knowledge. Against this, I argue that memory's constructive nature actually helps us to understand memory as a source, and remembering a species, of knowledge. I provide a positive account of remembering facts, objects, and events. In light of this account, I offer philosophical insights concerning memory's relation to other epistemic sources.

Table of Contents

Chapter 1 Varieties of Memory

This chapter introduces some background material to frame the project. In section 1, I discuss the notion of a 'memory system' as it is used in the empirical sciences of memory. In section 2, I distinguish memory systems from remembering states. Appreciating this distinction avoids conceptual confusion that has troubled some recent philosophical discussion of memory and motivates a taxonomic framework for the investigation. In section 3, I discuss theories of remembering. I begin by identifying two axes on which theories can vary: (i) preservative vs. generative, and (ii) epistemic vs. non-epistemic. I then describe four corresponding candidate theories. This sets the stage for chapters 2 and 3, in which I make a two-part argument for a generative epistemic theory of remembering.

1.1 Memory Systems

In contemporary psychology, a 'memory system' is a theoretical construct used to refer to a set of systems that enable behavior (broadly construed) to change as a function of experience (Johnson, 2007: 353).[1] Such systems are assumed to be instantiated in biological mechanisms, and so are taken to be collections of interacting brain areas that have a significant, though sometimes non-exclusive, role in memory functions.

Theorists distinguish memory systems from one another, and other psychological systems, by identifying dissociations using some combination of biological and functional criteria (Buckner 2007: 361-2). For example, Milner (1962) famously found that some severely amnesic patients could acquire and

[1] This is by no means to say that memory systems are the *only* systems that enable changes in behavior.

retain certain perceptual-motor skills (e.g. tracing the outline of a star in a mirror or learning to read mirror-reversed print) in spite of being unable to remember any personal experiences or any new conceptual information (i.e. the patients would have no conscious recollection of the tracing actions and no recall of the words read). This provided some motivation for recognizing two distinct memory systems: a "non-declarative" memory system that governs (among other things) the acquisition and retention of perceptual-motor skills, and a "declarative" memory system that governs the acquisition and retention of conceptual information and personal experience.

Another valuable observation involved a patient who, after suffering brain damage in a motorcycle accident, entirely lacked the ability to remember personally experienced events. He could not bring to mind any events or experiences from any period of his life. He did, however, retain much of the general knowledge he had acquired during his life. For example, he could read, write, play chess, and generally problem solve. He could also remember who he was, where he lived, where he had gone to school, the location of his family's summer cottage (on a map), and numerous facts about mathematics, geometry, and history.[2] This kind of dissociation partially motivated the recognition of two distinct declarative memory systems: one for general knowledge about the world ("semantic memory" *cf*. Quillian, 1968), and one for memories of life experiences ("episodic memory" *cf*. Tulving, 1972).[3]

[2] This research was pioneered by Nielsen (1958) and later Tulving (1988, 1985, 1972). The particular patient described here could also learn some new facts, though with considerable difficulty. Some take this to suggest that the boundaries between different memory systems are vague and involve considerable overlap. See Tulving (1999: 24-5) for discussion.

[3] See e.g. Squire (2007), Tulving (2000), Schacter and Tulving (1994), Eichenbaum (1993), Tulving (1983), Cohen and Squire (1980), Kolers (1975), and Scheffler (1965), among others, for more

The idea that memory can be roughly divided into distinct systems by using a combination of variables is now widely accepted.[4] However, there is no *generally* accepted taxonomy of memory systems, and there is a significant obstacle to providing one. Namely, many variables and criteria can be (and have been) used to dissociate systems from one another in many different ways, depending on one's theoretical interests. Consequently, any memory system classification scheme will focus, "rather arbitrarily on a few…differences while ignoring many others" (Roediger et al. 2002: 3). Below, I focus on four criteria that often frame philosophical discussions of memory.

1.1.1 Criteria

1) Duration: Memory systems are often distinguished (in part) in terms of duration. The behavioral effects of experience can last anywhere from a few seconds to indefinitely. Memory systems are said to be "short term" systems when they sustain behavioral effects for a few seconds or less. Memory systems are said to be "long term" when they sustain behavioral effects longer than just a few seconds – often indefinitely.[5]

2) Expression: Memory systems are also partially distinguishable in terms of the ways in which their behavioral effects are expressed. When they enable

---

discussion. It is also worth noting that many of the distinctions motivated by empirical observation had already appeared in one form or another in Bergson (1896), Cleparede (1911), Gall (1835), James (1890), Maine de Biran (1929/1804), Russell (1921), and Ryle (1949).

[4] See e.g. Hasselmo (2012), Buckner (2007: 361), Squire (2007, 2004), Schacter *et al*. 2000, Schacter and Tulving (1994), Schacter (1990), Roediger *et al.* (1990), and Tulving (1983) for discussion. It is worth noting, however, that the idea that memory can be divided into distinct systems is not *universally* accepted. See e.g. Surprenant and Neath (2009), Craik (2001), Howe (2000), Toth and Hung (1999), Weldon (1999), McClelland *et al*. (1995), and Roediger *et al.* (1990) for criticism of the "systems approach" to memory. Tulving (2002: 9), and Schacter *et al.* (2000: 630) suggest that disagreements on this issue are not particularly deep disagreements. I agree; moreover, I am inclined to think that this debate bears little on the philosophical issues of interest here.

[5] These boundaries are generally thought to be vague (cf. Roediger et al. (2002: 1).

behavioral changes that are (or at least can be) consciously and linguistically expressed by the subject (typically using declarative sentences), the governing memory systems are called 'declarative.' So, for example, remembering that Paris is the capital of France is governed by one or more declarative memory systems. In contrast, when experience gives rise to changes that are expressed via non-linguistic behavior only, e.g. via skillful performance, the governing memory systems are called 'non-declarative.' Swimming, playing the piano, and priming effects all depend on one or more non-declarative memory systems.

3) Information: Memory systems are also distinguished (partly) in terms of the kinds of information they process. Some memory systems represent states of affairs and events and so can be assessed for truth and/or accuracy. For example, memory for general facts about the world depends on what are often called 'semantic' memory systems; and memory for particular events in one's life depends on what are often called an 'episodic' memory systems. Other systems are "non-representational" in that they do not (or at least do not appear to) represent states of affairs or events. Rather, they encode procedural information, e.g. the information needed to successfully tie one's shoe or play the piano. Such abilities are said to depend on 'procedural' memory systems.

4) Biology: Finally, memory systems are also partially individuated by the biological (neurological) systems in which they are instantiated. Damage to different brain regions (specifically the hippocampus, medial temporal lobe diencephalon, striatum, neocortex, amygdala, or cerebellum) impacts memory capacities in different ways. So, for example, damage occurring in the striatum

> compromises one's ability to perform a variety of skillful non-linguistic tasks while leaving one's general knowledge about the world relatively in tact. Similarly, damage to the hippocampus compromises one's ability to recall facts and episodes while leaving one's perceptual-motor skills relatively in tact (Squire, 2007, Johnson 2007: 353, Roediger et al. 2002).

Philosophical interest has focused primarily on long term behavioral changes and this project is no different.[6] Next I summarize a simplified (but generally uncontroversial) taxonomy of long term memory systems.

1.1.2 A Basic Taxonomy of Memory Systems

Philosophers have focused primarily on issues related to declarative memory systems and I will follow this trend.[7] The two systems of most interest are characterized as follows. (i) A *semantic memory system* is a long term declarative memory system. It is distinctive in that it underlies the acquisition and retention of general knowledge (i.e. facts) about the world. Though said acquisition is (typically) causally dependent on prior experiences, such memory systems do not 'informationally' tie it to any specific experiences, i.e. it need not contain any information that explicitly refers to one's experiences (Tulving, 1983: 21). (ii) An *episodic memory system* is also a long term declarative memory system. It is

[6] I set aside all issues pertaining to short term memory systems such as the "working memory system" (WMS)--the system which actively holds information in the mind to do verbal and nonverbal tasks such as reasoning and comprehension, and which makes information available for further processing (see Becker, J. T., Morris, R. G. (1999) for some discussion). I will also set aside all issues related to the "Perceptual Representation System" (PRS) (See e.g. Schacter et. al. 2000) and Schacter and Tulving (1994: chapter 1)). This system is thought to be responsible for things like word and object *recognition* and is particularly valuable as part of an explanation of certain priming effects cf. Schacter (1990)).

[7] Philosophers sometimes do address one non-declarative memory system – the procedural memory system that governs perceptual-motor skills. Considerations of space prevent this category from playing any significant role in this project.

distinctive in that it operates on the domain of specific personal experiences, i.e. episodic memory systems *do* necessarily provide informational ties to one's prior experiences – they govern representations that involve explicit reference to one's prior experiences. In short, episodic memory is responsible for enabling us to "mentally travel through time" to "re-experience" past events in our lives (Tulving, Ibid).

While memory systems are primarily the domain of psychology and neuroscience, philosophers are also interested in the *products* of those systems – in the mental states one is in when one remembers. Nevertheless, they have not always been careful to distinguish the two and this has led to considerable confusion.[8] In the next section, I argue that the relationship between the taxonomy of memory systems and a taxonomy of remembering states that will be useful for this investigation is less straightforward than one might have expected it to be.

1.2 Remembering States

1.2.1 Criteria

Two sets of criteria have informed recent discussions of the taxonomy of remembering states: (i) the taxonomy of memory systems, (ii) the syntax of remembering reports.

1. Memory Systems: One *prima facie* plausible taxonomic scheme is guided the taxonomy of memory systems itself (or at least by the same considerations that govern that taxonomy). Roughly, if a remembering state is generated by an

[8] This isn't really unique to philosophers. Psychologists have blurred the distinction as well. See Russell and Hanna (2012), Sutton (2010), Schechtman (2011), Michaelian (2011), Hasselmo (2012), Hoerl (2007, 2001), and Tulving (2002) for some examples of instances in which the distinction between memory system and remembering state seems to have been inadequately appreciated.

episodic memory system, then it is an episodic remembering state; if it is generated by a semantic memory system, then it is a semantic remembering state; and so forth.

2. Syntax: Sven Bernecker (2010: 19-23) has proposed an alternative classification scheme in terms of the syntactic features of remembering reports. Specifically, his taxonomic system posits a distinct category of remembering for every possible combination of the verb "to remember" (or its near synonyms) and grammatical complement.

I argue that both sets of criteria yield taxonomies that are inadequate for an investigation into the nature of remembering.

1.2.2 Episodic and Semantic Remembering

Many have been tempted by a taxonomy of remembering states that parallels the taxonomy of memory systems I described in section 1.[9] However, it is not entirely clear why we should categorize remembering states in this way. As psychologists understand them, memory systems are large, complex, organized structures of more elementary operating components. They have fuzzy boundaries, share some components, and often *operate together to produce memories*. For example, the episodic and semantic memory systems often co-contribute to one's remembering general facts about the world and to one's remembering particular

[9] This taxonomy is extremely popular in the literature. See Shanton (2011), Sutton (2011), and Schechtman (2011) for recent endorsements of this strategy. It is worth noting that terminological labels vary in the literature. Semantic remembering is sometimes called "propositional memory" (Furlong (1951), Broad (1925), Bernecker (2010)) or "self-contained memory" cf. Matthen (2010). Episodic remembering is sometimes called "recollective memory", "personal memory", "experiential memory" or "direct memory" (cf. Bergson (1910), Russell (1921) Sutton (2011)).

experiences in their past.[10] Consequently, there is little reason to expect a one-to-one correspondence between memory systems and the types of remembering states they collectively produce.[11] Moreover, careful examination of the taxonomic categories shows that they fail to capture a variety of cases of remembering in any informative way. To see this, I will spell out the categories in detail and then present two example problem cases.

In the systems based taxonomy, *semantic remembering states* are characterized as follows: (i) they are typically reported using a that-clause as in 'S remembers that p', (ii) they involve conceptual information, (iii) they can occur without making reference to the events that are responsible for the subject's acquisition of the remembered information, and (iv) they need involve no imagery or characteristic phenomenology.[12] So, for example, my remembering that Cicero was a Roman orator, that Istanbul was once Constantinople, and that dogs have three eyelids, are all examples of semantic remembering.

In contrast, *episodic remembering states* are characterized as follows (cf. Tulving, 1972). (i) They are typically reported with a direct object as in 'S remembers the F' or a gerundive complement as in 'S remembers x'ing' or 'S remembers [x] φ'ing'.[13] (ii) They represent specific events occurring at a specific place and time. (iii) They are characterized by "mental travel in subjective time"—a

[10] See Schacter and Tulving (1994: 18), and Tulving (2002, 1991, 1983: 77-78).
[11] Bernecker (2010: 16-19) makes a related point. He argues that since this framework uses different sets of criteria to evaluate different remembering state types, it yields an inherently vague set of taxonomic boundaries. I'm inclined to think that vagueness itself is not the fundamental problem.
[12] See e.g. Hasselmo (2012), Bernecker (2010:15), Sutton (2004/2010), Tulving (1972: 386, 2002), Bower (2000), and Schacter et al. (2000).
[13] cf. Martin (2001). See also Sutton (2011) and Russell and Hanna (2012) for use of the gerundive construction.

distinctive re-experiencing of past episodes in which the rememberer has participated.

Though it is not clear that "mental time travel" and "re-experiencing" are ideal metaphors, it is clear that this aspect of episodic remembering has three defining features: (i) it is autobiographical, (ii) it involves imagery of the past episodes, and (iii) it is recognized, by the rememberer, to be about episodes in her past.[14] So, for example, when I remember the party last week, I 're-experience' the very event in my own past that is the source of my remembering by calling imagery of it to mind, and I now recognize that event as being in my past.

Despite its popularity, there are at least two familiar kinds of remembering that this framework does not adequately capture: (i) cases in which subjects do not recognize remembered events as in their past, and (ii) cases in which subjects remember events that are not part of their personal past. I begin with the latter.

Non-autobiographical event remembering

Remembering is *episodic* only if the remembered event/episode was an event/episode in the remembering subject's life. The justification for such a necessary condition is the intuition that it is impossible to remember an event that you did not witness or in which you did not participate. This is not quite right. Consider the following.

[14] See e.g. Sutton (2004/2010); Bernecker (2010: 13, 14); Squire and Kandel (1999: 106); Brewer (1996); Tulving (2002, 1983); Hoerl (1999: 235); Campbell (1994, 1997); and Schacter *et al*. (2000: 628). The term "mental time travel" was introduced by Tulving (1972); however, the idea that episodic remembering involves a distinctive kind of awareness of the remembered events is much older. See e.g. Russell (1921: chapters 9 and 12), James (1890), Locke (1690), and Aristotle (2001).

I sometimes call to mind images of President Obama standing on a platform in the center of thousands of people, of him embracing the First Lady, of him smiling while being sworn in, and so forth. When I do so, I either remember Barack Obama's 2009 inauguration or his 2012 inauguration. However, like many others, I did not attend either of these inaugurations, nor did I watch them in live time. Rather, I watched tape-delayed coverage of those events and it was only by doing so that I can now remember as I do.[15]

Of course, just as I can remember relatively recent events like Obama's inaugurations, I can also remember events from the more distant past. For example, I can call to mind images of the Challenger sitting on the launch-pad, lifting off into the air, and exploding shortly thereafter. When I do so, I remember the Challenger Disaster. And just as with Obama's inauguration, I (like many others) can only remember the disaster as I do in virtue of having watched delayed coverage of it.

In such cases, one remembers events. At the very least, it would be implausible to maintain that these are instances of *semantic* remembering.[16] Although I do come to know facts about the world in virtue of remembering as I do, e.g. that Obama stood on a platform and that the Challenger exploded shortly after take-off, I do so only *indirectly* in virtue of being in remembering states that essentially involve imagery of past events in my life and that I recognize as such.

Compare these cases with others in which someone remembers the aforementioned facts simply in virtue of having received testimony about them. In such cases, one need have no imagery of the corresponding events; imagery is

[15] See Jacobson (2005: 274) for a similar point.
[16] Bernecker (2010: chapter 1) would disagree.

inessential to semantic remembering. It is not, however, inessential to the cases described above; rather, the imagery is partly constitutive of my remembering, and as such, I have the propositional knowledge described above in virtue of undergoing states that are partly constituted by that imagery. This suggests that even if I can truly be reported as remembering certain facts, I am not *merely* semantically remembering. At most, I am semantically remembering in virtue episodically remembering.[17]

Moreover, just as one can see an event e, without seeing *that* e is occurring, I could very well remember as described above without going on to deploy any of the concepts that are necessary for remembering the associated facts. For example, in the inauguration case, I might be entirely unaware of the fact that I'm remembering the inauguration at all; I might wonder (or even have no opinion as to) whether the imagery I call to mind is imagery of his inauguration, a campaign event, or his 2004 speech at the Democratic National Convention. None of this challenges the idea that in calling to mind images of the inauguration, as I did, I come to remember *it*.[18]

Once we see the structure of this kind of case it is clear that the same kinds of mechanisms could enable me to remember events from still longer ago. For example, I might remember Hinckley's attempted assassination of Reagan or JFK's actual assassination. This presents a problem for the category of episodic remembering as it is generally defined because there is no sense in which these

[17] Bernecker (2010: 18) considers but dismisses this possibility as implausible. I don't see why such a possibility would be implausible in this kind of scenario. Indeed, it seems to me that this is the right thing to say.

[18] This argument could be made equally effectively by focusing on 'his standing' if one is worried that such imagery may not suffice (in conjunction related causal and epistemic connections) to make the remembering about the larger inauguration event

events are prior events in my life – I was not alive when they occurred. Consequently, such remembering cannot meet the autobiographical condition, even in its most minimal form. The upshot is that these instances of remembering are not satisfactorily categorized as either instances of semantic remembering or as instances of episodic remembering. So, the memory systems based taxonomic framework leaves familiar examples of remembering unclassified.

One might try to resist this conclusion by offering an alternative characterization of my remembering states. Specifically, one might argue that although I am episodically remembering in each case, I'm not actually remembering Obama's inauguration, the Challenger Disaster, or the assassination events. Rather, I'm remembering prior experiences of mine; namely, my experiences of watching tape-delayed recordings of those events.

This assessment of the cases requires some bullet-biting. For example, it has the consequence that very few people remember Obama's inauguration—even if they can call to mind images of the event, report on it accurately, and would confidently report themselves as remembering it. It is hard to see why one should accept such consequences for the sake of preserving a classification scheme.

A better suggestion can be derived from Mike Martin's account of episodic remembering. On his view, *all* episodic remembering is remembering of our prior *experiences* of events.[19] One sympathetic to Martin's view might argue that I really

[19] Martin (2001). Strictly speaking, Martin introduces the technical term "apprehension" in place of experience. Although it is not clear that he takes the terms "experience" and "apprehension" to be interchangeable, he does not provide a definition of apprehension that would distinguish it from experience and so I will read him as claiming that to remember an event is to remember one's experience of it. See Von Leyden (1961) for a precursor to Martin's view.

do episodically remember all of the aforementioned events. I simply do so in virtue of episodically remembering my (mediated) experiences of them.

Although this proposal has some merits, I don't think that it is ultimately sustainable. First, it threatens to collapse a natural distinction that is often tracked with the following two constructions: (i) 'A remembers the F', (ii) 'A remembers x'ing.' The former kind of construction is naturally used when one remembers an object, property, or event, independently of reference to any relation between that event and the self. The latter construction is naturally used when one remembers an event and one's relation to it.

Some example cases should make the distinction clear. It is natural for me to say that I remember the *Mona Lisa*. However, given that I have never been to the Louvre, and that I don't recall any of the occasions upon which I have viewed any depictions of it, it would be odd for me to say that I remember seeing (or experiencing) the *Mona Lisa*. The fact that I only now remember it in virtue of having seen depictions of it at some time or other should not be confused with my remembering seeing those depictions of it. To see why, compare a case in which I do remember seeing something; for example, I remember seeing *American Gothic* during my trip to the Art Institute when I was last in Chicago. Remembering *seeing American Gothic* involves imagery not just of the piece itself, but also my perspective on it, the surrounding crowd, my emotional responses, and so forth. Remembering the *Mona Lisa* itself need involve none of those things.

Similarly, it is natural to say that I remember Obama's inauguration and that I only remember it in virtue of having seen recording(s) of it. It is, however, odd for

me to say that I remember seeing recordings of Obama's inauguration. In fact, I don't seem to remember seeing such recordings at all. That I must have seen recordings of it is something I'm aware of now only in virtue of inferring that fact on the basis of my background understanding of how one typically comes to encounter such events.

To maintain the view that remembering an episode is just remembering one's experience of that episode, one will need to provide some additional theoretical resources for collapsing this distinction. Whether Martin's position has such resources is unclear. In the above cases, one might try to explain away the distinction by saying that I am remembering my prior experiences in all of these cases, but I only conceptualize it *as* remembering my prior experience in the cases I have claimed are naturally reported with a gerund phrase. This, however, does not seem to do justice to the differences between these cases. Remembering seeing something involves more than just additional conceptualization; it requires additional information about the event of seeing itself (e.g. reference to oneself, imagery from one's own perspective, and or additional contextual information).

Finally, one could also try to motivate Martin's position with an appeal to transparency. It is plausible that one's perceptual experience of an episode is transparent; what it's like to undergo a particular (visual) experience of the episode is just the way the episode is represented as being.[20] If we assume that memory experiences are also transparent, holding that one remembers an episode in virtue of remembering her experience of it looks somewhat plausible. The basic idea would be that transparency is a transitive property. The memory experience is

[20] See e.g. Tye (forthcoming) for more on transparency

transparent to the prior perceptual experience and, in virtue of perceptual experience being transparent to the perceived event, the memory experience is transparent to that event as well.

Unfortunately, it is not obvious that transparency is transitive in this way. Perceiving and remembering appear to have some crucial differences. One particularly interesting difference concerns perspective. Perceptual experiences have a fixed perspective. One sees an event from the 1st-person point of view and though she may change her orientation during a particular episode, her perception of it remains from the inside. As is now well known, this is not always so for remembering. Subjects can remember events from the 1st-person perspective ("field perspective"). They can also sometimes remember the *very same event* from an "external" perspective ("observer perspective"), i.e. from a perspective in which they have an image of themselves as a participant in the event.[21]

This does not refute a transparency thesis for memory experience but it does have interesting consequences for its shape in memory contexts. Specifically, it seems to show that Martin's *prima facie* plausible view that one remembers an episode in virtue of remembering her experience of it is at least sometimes false. When one remembers an episode from the "observer perspective" the episode is represented as having at least some properties that her prior experience of the

[21] Some have reported switching perspectives in the course of one remembering episode—at one moment remembering an event from the inside and at the next, remembering it from an external position. See e.g. Hasselmo (2012), Matthen (2010), Sutton (2010), Conway (2009), Robinson and Swanson (1993), Peacocke (1985), Nigro and Neisser (1983), Williams (1973) for discussion.

event did not have (assuming that that experience was itself transparent): namely, the memory experience includes visible properties of the self.[22]

Given this, a natural conclusion is that one can remember an event, and that her memory experience can be transparent to that event without remembering her experience of the event (and without the memory experience picking up its transparency from the transparency of the prior perceptual experience). This is not to say that remembering is not causally dependent on prior perceptual experiences; it clearly is. Moreover, the memory representation may even be partially constituted by representational information preserved from the prior experience. It simply doesn't follow from this that all remembering of events is remembering of one's prior experiences. Thus, formulations of episodic remembering that involve even a minimal autobiographical condition will exclude a variety of cases of remembering that don't naturally fit elsewhere in the taxonomic scheme being considered here.

Remembering without recognition

A second problem for the memory systems based taxonomy involves cases in which subjects remember objects or events without those objects and events seeming to them to be in the past. Consider the following example:

> Suppose that someone asks a painter to paint an imaginary scene. The painter agrees to do this and, taking himself to be painting some purely imaginary scene, paints a detailed picture of a farmyard including a certain colored and shaped house, various people with detailed features, particular items of clothing and so on. His parents then recognize the picture as a very accurate representation of a scene which the painter saw just once in his childhood…the painter did his work by no mere accident. (Martin and Deutscher, 1966: 167-8)

[22] One could potentially devise a scenario involving mirrors in which an image of the self is present in the visual experience, but such possibilities do not diminish the point being made here.

Martin and Deutscher conclude that although the painter sincerely believes that his work is purely imaginary he has actually remembered a scene from his childhood and subsequently painted it. Importantly, in such a situation he remembers the scene even though he has no awareness of it "coming back" to him and does not judge that he is remembering or that the scene is one from his past.

Since the painter fails to recognize the scene as one he had seen, he cannot be episodically remembering as it is normally defined. Nevertheless, such remembering is tied directly to one particular event and it crucially involves imagery of that event. This would suggest that it pairs better with episodic remembering than it does with semantic remembering. Moreover, it is implausible to say that the painter merely remembers *that* there were such and such trees, animals, clothing, and so forth. Indeed, it doesn't seem that he remembers *that* there were such things at all. Having intended to paint a fictional scene, he would not have painted them if he had remembered those facts. So, since some remembering of past events and objects can occur without any judgment that it has occurred in the subject's past, formulations of the category of episodic remembering that require a recognition condition will rule out a variety of cases of remembering that don't naturally fit elsewhere in the basic taxonomy.[23]

The upshot is that the memory systems-based taxonomic framework is overly restrictive. An investigation into the nature of remembering needs a taxonomic framework that recognizes more than just the remembering of facts and the remembering of episodes in one's life.

[23] See chapter 4 for additional discussion of related cases.

1.2.3 The Syntactic Taxonomy

Individuating remembering state types in terms of the syntactic features of remembering reports, as Bernecker (2010) does, yields the following taxonomic categories.

1. *Propositional Remembering* refers to remembering states that are reported with any substituend of the schema 'S remembers that p'. So, 'Ali remembers that Simon likes cheddar', and 'Jill remembers that she grew up in Baltimore' are instances of propositional remembering.
2. *Direct Object Remembering* refers to remembering states that are reported with any substituend of the schema 'S remembers x' where 'x' is a noun phrase. Accordingly, 'John remembers Jane', 'Jeremy remembers the 4th of July', and 'Jess remembers her favorite Beatle' are all instances of direct object remembering.
3. *Gerundival Remembering* refers to remembering states that are reported with any substituend of the schema 'S remembers x'ing' where 'x'ing' is a gerundival complement. So, when I remember going to the store or hearing the alarm, I am engaged in gerundival remembering.
4. *Infinitival Remembering* refers to remembering states that are reported with any substituend of the schema 'S remembers to x' where 'x' is a verb, e.g. 'Jerry remembers to lock the door whenever he leaves.'
5. *Interrogative Remembering* refers to remembering states that are reported with any substituend of the schema 'S remembers wh-phrase' where the relevant 'wh' phrases are interrogative pronouns such as 'who', 'where', 'what', 'when', or

'why'. Sam remembers who stole the diamonds, what direction they ran in, when the police arrived, etc. are all instances of interrogative remembering.

6. Finally, *How-Remembering* refers to remembering states that are reported with constructions of the following form, 'S remembers how…' as in 'John remembers how to play the piano' or 'Jane remembers how Jim broke the vase'.

This taxonomic schema is an improvement in so far as it avoids the problems raised for the memory systems based taxonomy by simply omitting autobiographical and recognitional requirements from any taxonomic categories. However, it is also not adequate for an investigation into the nature of remembering because, like other attitudinal reports, remembering reports only provide partial guidance to the features of one's mental states. In particular, the grammatical features of remembering reports fail to provide much information about the objects of remembering states. The problem is that, in many cases, syntactically distinct reports can be properly used to report the very same remembering states, and a single report form can be used to report states with distinct objects. To see this consider the following examples.

First, a direct object construction can be used to truly report instances when which subjects remember objects or events, as well as instances in which subjects remember facts involving those objects or events. For example, 'Lana remembers the capital of Spain' can be true if she answers "Madrid" when asked what the capital of Spain is. It can also be true when she calls to mind images of the different parts of Madrid that she visited during her most recent trip to Spain. Note that, in the latter case, it would remain true that Lana remembers the capital of Spain even

if she does not know, and would not report, *that* Madrid is the capital of Spain. Similarly, reports like "Jim remembers how to bake a brie" can be true in contexts in which Jim can describe a list of instructions for doing so, or in contexts in which, despite being unable to list the instructions, Jim successfully bakes a brie.

In the other direction, reports like 'Woodhouse remembers who shot Reagan' and 'Woodhouse remembers that Hinckley shot Reagan' appear to report the very same phenomenon. The truth of both reports seems to depend on the very same things as, in either case, the report of Woodhouse's remembering takes a particular fact as its object, i.e. the fact that Hinckley shot Reagan. These kinds of examples, and further possible permutations, strongly suggest that developing a taxonomy of remembering states in terms of the syntactic features of remembering reports tells us too little to be particularly useful for an investigation into the nature of remembering. Something else is needed.

1.2.4 A Semantic Taxonomy

A more promising option, and the one I will adopt for this project, distinguishes taxonomic categories of remembering states, not in terms of the syntactic features of remembering reports, but rather in terms of their semantic objects, i.e. the different kinds of metaphysical entities remembering states can be about.[24] A taxonomy developed in this way will have as many categories as there are distinct possible objects of remembering. However, how many an investigation

[24] Though Bernecker (2010:19) explicitly commits to using a syntactic criterion to individuate remembering state types, his subsequent discussion sometimes suggests that he is at least sympathetic to a semantic criterion.

chooses to recognize will differ depending on one's theoretical interests. I focus on five categories below.

1) *Facts Remembering* refers to those remembering states that take facts as their objects. Such states are often reported using propositional constructions as in 'S remembers that p' but they can be reported using direct object, remembers-wh and remembers-how constructions as well.
2) *Object Remembering* refers to those remembering states that take concrete particulars as their object, e.g. persons, places, or other concreta, and is typically reported using direct object constructions.
3) *Event Remembering* refers to those remembering states that take events as their objects. They are typically reported using direct object constructions where the embedded noun depicts an event, as in 'S remembers the party', or using gerund phrase constructions as in 'S remember going to the party'
4) *Property Remembering* refers to those remembering states that take properties as their object. They are typically reported using direct object constructions as in, 'S remembers the color of the getaway car'.
5) *Mental State Remembering* refers to those remembering states that take mental states as their objects. They are typically reported using direct object, 'wh' or 'how' constructions as in, 'S remembers how she felt about the election,' or 'S remembers what he thought when he heard the news,' or 'S remembers the joy she felt when she won the award'.

Like the syntactic taxonomy, this semantic taxonomy of remembering states refrains from providing any autobiographical or recognitional constraints on

remembering and can categorize the kinds of cases that troubled the memory systems-based taxonomy. However, it is also an improvement over a syntactic taxonomy because it respects the variability of the relationship between remembering reports and remembering states. Consequently, the semantic taxonomy puts us in a good position to begin an investigation into the nature of remembering.

Before proceeding, I end this section by addressing a potential worry. I argued that an adequate taxonomy of remembering states cannot include recognitional or autobiographical conditions as necessary conditions on remembering state types. It does not follow from this such features of remembering states are irrelevant, and one might worry that a taxonomic framework that forces us to ignore such features has deep problems of its own.[25]

Fortunately, the semantic taxonomy I have sketched here does not have this consequence. Some memories certainly are autobiographical. For example, while one can remember an assassination without witnessing it directly, one cannot remember *witnessing* said assassination if one did not witness it. This distinction is not captured by some special necessary condition on remembering state types. Rather, it is captured by the factivity of remembering. In this case one cannot remember witnessing it because it is not true that one witnessed it. Nevertheless, given that the event did take place, being unable to remember witnessing it is no, in principle, obstacle to being able to remember the event itself.

[25] See e.g. Sutton (2011) and Schechtman (2011) for some criticism of Bernecker's 2010 on these grounds.

Similarly, we do often recognize things we remember as being in the past. While I argued that such recognition cannot be a necessary condition on a remembering state type, it is nevertheless a valuable feature of certain remembering states in that it often allows a subject, not only to remember the event, but to come to know that they are remembering the event.

Having described the semantic taxonomy of remembering states that will guide this project, I conclude this chapter by characterizing possible types of theories of remembering.

1.3 Theories

Theories of remembering vary on two central axes: (i) they can be preservative or generative, and (ii) they can be epistemic or non-epistemic. According to preservative theories, the faculty of memory has only the capacity to preserve features that have been generated by other purportedly more basic sources, such as perception and reason (Lackey, 2007: 209 fn. 2).[26] In contrast, generative theories allow that memory can generate things, e.g. representations, beliefs, justification, and knowledge.[27]

---

[26] See, e.g. Audi (1997: 410), Dummett (1993: 420-1), Plantinga (1993: 61 n22), Malcolm (1963: 223; 1977: 102-8).

[27] See e.g. Bernecker (2010), Lackey (2005), and Michaelian (2011). Two points of clarification: (i) Some (e.g. Senor, 2007) draw a further distinction between theories that are preservative (or generative) with respect to *epistemic* features (justification or warrant) and theories that are preservative (or generative) with respect to the *content* (or content-vehicle) of the remembering state. As I will ultimately advocate a view that is generative in all of these respects, I will focus only on the broader distinction. (ii) Some (e.g. Pollock, 1974: 193) hold that the phenomenology of remembering is epistemically significant and so remembering is, in a sense, *always* generative. As discussed in section two, the phenomenology of remembering states can vary and be absent or misleading. Consequently, generative theories, as I will understand them, do not carry such a commitment (see also McGrath 2007: 19-22 and Bernecker 2010: 96-100 for some critical discussion).

According to epistemic theories, remembering is fundamentally a kind of epistemic success, e.g. knowledge.[28] In contrast, non-epistemic theories do not carry such a commitment. They allow that remembering *can* give rise to one or more kinds of epistemic success, but they deny that it is itself an epistemic success state and they maintain that remembering need not contribute to any further epistemic success state either.[29] Taken together, these axes give rise to four possible kinds of theories: *Preservative Epistemic Theories*, *Preservative Non-Epistemic Theories*, *Generative Epistemic Theories*, and *Generative Non-Epistemic Theories*. I discuss them each in broad strokes before arguing for a version of generative epistemic theory in chapters two and three.

1.3.1 Preservative Epistemic Theories

A preservative epistemic theory combines two central thoughts. (i) Memory is fundamentally preservative; it cannot be a source of new information or knowledge; rather, at best, it preserves it from prior sources. (ii) Remembering is fundamentally an epistemic success state.

In chapter two, I will examine epistemic theories in more detail, but the paradigmatic preservative view treats remembering as a kind of knowing. So, for example, 'if you remember that we met, you know that we did. Similarly, if you remember me, you know me' (Audi, 2003: 69). In sum, a preservative epistemic theory roughly holds that, 'To remember now is to know now what you knew in the

[28] See, e.g. Annis (1980: 324), Anscombe (1981: 127), Ayer (1956: 138), Dancy (1985: 187), Dretske (1981: 361), Evans (1982: 235), Grice (1941: 344), Moon (2013), Pollock (1974: 196), Ryle (1949: 272-9), Shoemaker (1970), Squires (1969: 185), Unger (1972: 304), Williams (1973: 142), Williamson (2000: 37-8), and Zemach (1968: 529).
[29] See e.g. Bernecker (2011, 2010, 2008, 2007), De Brigard (2013, 2012), Lehrer and Richard (1975), Michaelian (2013, 2011), Sutton (2010, 2009).

past, without learning in-between what you know now…Memory…is *knowing from the past*' (Margalit, 2002: 14).

1.3.2 Preservative Non-Epistemic Theories

A preservative non-epistemic theory retains the idea that memory is fundamentally preservative, but drops the idea that it need be a kind of epistemic success. So, such a theory holds that, if you remember that we met, you at least, believe or represent that we met, and your present representation is preserved from your past representation that we met. Similarly, if you remember me, you represent me and your representation comes from your prior representation of me.[30]

In sum, a preservative non-epistemic theory of remembering holds roughly that to remember now is to believe, or perhaps represent, now what you believed or represented in the past, without re-acquiring in-between what you believe or represent now. Of course, the most plausible way to develop such a theory, without appealing to any essentially epistemic features of remembering, is in terms of the causal relations between one's present representations and one's prior representations, and I will focus subsequent discussion on causal theories.

1.3.3 Generative Epistemic Theories

A generative epistemic theory of remembering retains the idea that remembering is fundamentally an epistemic success state, but drops the idea that memory has *only* the capacity to preserve features that have been generated by

[30] Characterizing the view in this way assumes a broadly representational theory of mind; however, nothing turns on this. One drawn to relational or naïve realist theories of mind could adjust the formulation accordingly.

other sources. Such a theory does not require that memory is exclusively generative, i.e. it does not claim that *nothing* is preserved by memory, nor does it claim that memory *always* generates novel knowledge; it requires only that it *can* generate novel knowledge in at least some circumstances.

So, according to such a theory, if you remember that we met, you know that we did, and if you remember me, you know me. However, it allows that one might also come to know something new by remembering it. For example, one might come to know that one was a sensitive child by remembering that one was a sensitive child, without having ever before represented that fact to oneself. In sum, according to a paradigmatically generative epistemic theory, to remember something is to know it by operations of memory, regardless of whether one previously knew it.

1.3.4 Generative Non-Epistemic Theories

Finally, a generative non-epistemic theory of remembering drops both the idea that memory is fundamentally preservative, and the idea that remembering is fundamentally a kind of epistemic success. According to such a theory, if you remember something, you at least represent it, though you need not know it, and need not have known or represented it at any prior time. And as with a preservative non-epistemic theory of remembering, the most plausible way to develop such a theory is in terms of the causal relations between one's present representations and one's prior representations.

So, if you remember that we met, you represent that we met, though you need not know that we did, nor need you have previously represented that we did.

Your present representation must be caused, in some to be specified way, by your prior representations and mental states. The central challenge for a generative non-epistemic theory of remembering is to develop these ideas in a compelling way, and recent attempts to do so are prominent in the philosophical literature (see, e.g. Bernecker, 2010 and Michaelian, 2011).

1.4 Conclusions

In this chapter I distinguished the notion of a memory system from the more philosophically interesting notion of a remembering state. I then argued against modeling one's taxonomy of remembering states on a taxonomy of memory systems. I further argued against using the syntactic features of remembering reports to individuate remembering state types. In lieu of these two taxonomic schemas, I adopted a semantic taxonomy of remembering states developed in terms of the kinds of metaphysical entities remembering states can be about. Finally, I described four possible kinds of theories of remembering. In chapter two, I argue for a generative epistemic theory of remembering by way of an argument by elimination. In chapter three, I provide some positive arguments for a particular version of epistemic theory of remembering, and develop the view for the remembering of facts.

Chapter 2 Paths Not Taken

This chapter provides the first of a two-part argument for a generative epistemic theory of remembering. It has the form of an argument by elimination. At the end of chapter one, I outlined four basic kinds of theories of remembering. In this chapter, I argue that three of those four are ill suited to provide a compelling account of what it is to remember something.

In section 2.1, I argue that preservative theories (both epistemic and non-epistemic) are subject to plausible counterexample cases. Moreover, they are in tension with our best understanding of *how* human memory processes work. This motivates pursuit of generative theory of remembering. In section 2.2, I argue that generative non-epistemic theories of remembering (i.e. causal theories of memory) have deep problems of their own.[1] First, the motivations generally proffered on their behalf provide equal or better motivation for an epistemic alternative. Second, they fail to provide a set of sufficient conditions for remembering. Third, current formulations of those theories fail to provide a set of necessary conditions for remembering. Moreover, augmenting such theories appears to require the inclusion of epistemic features into the analysis. Consequently, some form of generative epistemic theory is the most plausible candidate for a compelling theory of remembering.

[1] One might question whether there are any interesting non-causal and non-epistemic theories of remembering. While I will not treat this issue at any length, it is difficult to see how such a theory might be developed. A memory state must have some kind of connection to a prior cognizing state. Such a connection is either explained or unexplained and inexplicable necessities look implausible here. The most plausible explanation of such a connection appeals to causation either as partially constitutive of memory or as a channel by which we explain something that is partially constitutive of memory (e.g. knowledge). So, it looks like any plausible theory of remembering will be causal or epistemic in nature.

2.1 Against Preservative Theories

Preservative theories of memory can be traced back at least to Plato's *Theaetetus,* where he likens memory to a block of wax in which perceptions are imprinted.[2] The central idea for any preservative theory is that the function of memory is to preserve whatever it receives from more basic informational sources. So, for example, if perception provides it with information *x*, about object *y,* that is justified to degree *z* at t_1, memory performs its function appropriately only to the extent that it retains *x* (and its epistemic status) and makes it available for use at some later time t_2.

Though this idea may initially appear tempting, it is not sustainable. In particular, it is subject to at least three plausible kinds of counterexample case: (i) cases in which one's relationship to evidence at t_2 is better than one's relationship to evidence was at t_1, (ii) cases in which memory processes *add information* between t_1 and t_2, and (iii) cases in which memory otherwise *transforms* representations between t_1 and t_2. I describe examples of each in section 2.1.1.

2.1.1 Counterexample cases

Improving the evidential relation

One problem for preservative theories of remembering is that, at least sometimes, one can be in a better epistemic position when remembering at t_2, than they were at t_1. Jennifer Lackey (2005) offers two such cases that appear to *generate novel* knowledge when the subject remembers.

[2] Augustine (1991: 191 (circa ce 397), Locke (1690: 149-50, 152-3) and Broadbent (1971) provide a similar analogy between memory and a "storehouse" or library. See Sutton (2010) for some additional discussion.

Case 1: Doxastic defeaters

> While an undergraduate in college, Nora was a very careful and epistemically reliable recipient of testimony, with one notable exception: she was overly susceptible to peer-pressure from two of her friends who belonged to a religious cult. After repeatedly hearing them rant and rave about the corrupt minds and souls of non-believers, she eventually became convinced that the testimony of atheists is nearly completely unreliable. During this time, Nora had several conversations with Calvin, a fellow student in one of her classes who, as a matter of fact, was an extremely reliable source of information and whom she had every reason to believe was both competent and sincere with respect to his reports. Yet Nora also knew that Calvin was an atheist, and so she believed him to be a highly unreliable epistemic source. One day after class, they were discussing World War II and Calvin told Nora, much to her surprise, that Hitler was raised a Christian. Being momentarily caught off guard, Nora found herself believing this proposition on the basis of Calvin's otherwise epistemically flawless testimony.
>
> Now, several years after graduating from college, Nora is no longer in touch with her friends who were members of the religious cult and she has ceased believing that the majority of the testimony offered by atheists is highly unreliable-such a belief has simply faded from her memory. At the same time, however, she still believes on the basis of memory dating back solely to Calvin's testimony that Hitler was raised a Christian. [3]

According to Lackey, this is an example in which someone knows something by remembering it, despite having not previously known it because memory has enabled her to improve her epistemic situation by jettisoning a defeater. The structure of this case is as follows. At t_1, *S* formed belief *b* on the basis of evidence *e*. At t_1, she did not know *b* for she also held belief *d* that served as a defeater for *e*. At t_2, *S* no longer held belief *d* and so no longer had a defeater for *e*. Further, her memory belief *b* at t_2 remained, as one would expect, based on *e*. So, at t_2, *S* had a true, justified, and undefeated, belief *b*. And so, she would standardly be thought to have known *b*.[4]

It is important to note that Lackey's case does not show that memory improves a remembered belief's epistemic status whenever it is remembered, or

[3] Lackey, 2005: 644-49. See also Bernecker, 2010: 100 for a structurally analogous case.
[4] Lackey, 2007: 210

whenever one forgets something inconsistent with a remembered belief.[5] It shows only that memory *can* support epistemic improvements in some circumstances and so *can* be a generative epistemic source.

In this case, adopting the defeating belief (i.e. the belief that the testimony of atheists, and so Calvin, is unreliable) is an unfortunate development for Nora. She is wrong about Calvin when forming such a belief. Nevertheless, once she has it, it prevents her from being able to know what she came to believe on the basis of his testimony. It is simply a fortunate development for her that memory did not retain the misleading belief. Its failure to do so allows the belief she formed on the basis of otherwise epistemically flawless testimony to achieve its epistemic potential – to be knowledge. An analogy with testimony helps to illuminate this point. Let us allow that, at the time of recall, retained background beliefs function a lot like testimony does at the time of initial belief acquisition. Just as one can be fortunate to stop receiving misleading information from interlocutors when forming beliefs, one can be fortunate to stop receiving misleading information from one's memory at the point of recall. In both cases, the elimination of certain information can itself improve one's epistemic position, enabling one to know that *p*. This possibility is incompatible with a preservative theory of remembering because memory does not merely preserve a prior belief's epistemic status. Rather, it supports an improvement in epistemic status at the time of remembering.

Case 2: Normative defeaters

[5] See Audi, 1995 and Pollock, 1974 for characterizations of memory that would have this commitment. See Bernecker, 2010: 96 and McGrath, 2007: 19-22 for critical discussion of such views.

A subject's epistemic position can arguably be improved when the world changes around her in the right sorts of ways as well. In particular, one can arguably come to know something that one did not previously know (by remembering it) when changing circumstances eliminate a normative defeater of which one was unaware. For example, consider the following case:

> On Tuesday, everyone knows that the king is dead. They know this because they all watched the funeral on television, read his obituary in the papers, and so forth. On Wednesday, the government undertakes a massive deception—telling everyone that the king did not really die; the previous "news" was all an elaborate hoax. They even support this claim by orchestrating a news conference at which the king "appears" to assure the people that reports of this death were greatly exaggerated. Suppose further that everyone but Clyde gets the message. They all now believe that the king is alive and well. Clyde, however, has been cloistered in his study working on his thesis since Tuesday night. As a result, on Thursday, he still believes that the king is dead, and he believes it on the same grounds that everyone (originally) had for believing this. Now, suppose that on Friday the cover-up conspiracy is revealed and the newspapers once again report that the king is dead. Again, everyone revises the beliefs accordingly, and again Clyde is unaware of any of these developments.[6]

Some (e.g. Bernecker and Lackey) interpret this kind of case as follows. On Thursday, no one knows that the king is dead, including Clyde. The general community does not know it because the cover-up has provided them with undefeated doxastic defeaters. Clyde doesn't know that the king is dead because he holds the belief only in virtue of having failed at his epistemic duties (i.e. in virtue of not keeping up with the news reports) – Clyde's true belief is subject to an undefeated normative defeater. On Friday, the normative defeater – the one thing preventing Clyde from knowing that the king is dead – has been defeated. In its absence, Clyde is once again able to know that the king is dead.

[6] This version of the case is adapted from Dretske and Yourgrau, 1983. See also Lackey, 2005: 640-3, Bernecker, 2010: 101-104 and Harman, 1973: 143-4 for similar cases.

Though I do not accept this interpretation of this particular case (see chapter 5 for more discussion), those that do will take it to provide even more evidence against preservative theories of remembering in virtue of the fact that it shows that a subject's relation to normative defeaters can change over time.

Additions

Counterexamples to preservative accounts of memory also come from the fact that memory can sometimes *add* to one's prior representations when one remembers. I will discuss two ways in which this can happen.[7]

Case 3: Adding attitudes

First, memory can sometimes generate a token of a new attitude type when one remembers – specifically, it can give rise to beliefs where there previously were none. Consider another of Lackey's cases (2005: 650):[8]

> Yesterday morning was like most others for Clifford: he spent it drinking coffee, listening to the radio, and driving in his car during his hour-and-a-half commute to work. As was typical for these commutes, Clifford's attention was divided between the other cars on the road, the surrounding environment, the discussion and music on the radio, and his thoughts about the day's work. Because of this perceptual and cognitive overload, Clifford found himself, as he often did on these drives, taking in more pieces of information than he actually processed at that time.
>
> Indeed, this was made apparent earlier this morning, when Clifford bumped into his friend, Phoebe, at the bakery and started talking about his commute. During this conversation, Phoebe asked him whether construction had begun on I-55. Though this is not the freeway that Clifford takes to work, he does pass it every day and, moreover, it is the route that he occasionally takes to a nearby shopping center. Upon being asked this question by Phoebe, Clifford paused, called to mind passing I-55 on his drive to work yesterday, and correctly remembered seeing construction work being done on this freeway. He, therefore, responded affirmatively to Phoebe's question, adding that he will be sure to map some alternate routes so as to avoid the traffic delays inevitably brought by construction. Prior to the recollection of the visual image triggered by this question, however, Clifford would have continued taking I-55 to the shopping center and wouldn't have made even minor efforts to

[7] These are not plausibly the *only* ways for memory to add things before/during remembering. See, for example, De Brigard, 2014; Michaelian, 2011; Park et al., 2007; and Intraub et al., 1992 for discussion of the phenomenon of "boundary extension" – cases in which subjects remember more of a visual scene than they initially perceived.

[8] See also Bernecker, 2010: 85 for a parallel case.

avoid this freeway.

Bernecker (2010: 99) points out that this is not a case in which memory has generated new justification for a belief. Given some assumptions about the reliability of his perceptual and cognitive systems, the information Clifford visually processed at t_1 was already fully justified. Nevertheless, this case remains a counterexample to preservative theories of memory (according to Lackey) because Phoebe's question prompted Clifford to recall a prior visual image of the traffic construction he saw the day before. Then, having done so, Clifford came to form a *new* belief, namely, that there is construction on I-55. So, contrary to a preservative theory, memory can give rise to new epistemic features – namely, attitudes.

One might be tempted to resist Lackey's conclusion by arguing that Clifford really did believe that there was construction on I-55 when he saw it – his belief was simply a dispositional one. However, Lackey (2007: 217) points out that there are some important differences between this case and ordinary examples of dispositional belief. Ordinarily, transferring a dispositional belief to an occurrent belief has relatively little impact for the subject. To use Lackey's own example, it is plausible that one dispositionally believes that oranges do not grow on kangaroos. Consciously entertaining that belief has little effect. It might, for example, lead one to say, "of course oranges don't grow on kangaroos", but it would not modify other subsequent behavior with respect to oranges and kangaroos because one *already* held the belief.

In contrast, everything about Clifford's behavior prior to Phoebe's prompting supports the conclusion that he had no beliefs about any construction on I-55.

Moreover, after the prompting and subsequent recall, Clifford actively did modify his behavior in light of his belief about the construction. This, we might think, best supports the conclusion that memory really has generated a novel attitudinal state. Though he may have perceptually represented the construction at t_1, Clifford came to *believe* (and know) something new by remembering that there was construction, at t_2.

Case 4: Conceptual enrichment

The third case was one in which perceptual information was put to use in the formation of a novel belief. Memory can also add information to beliefs themselves and, by doing so, enable one to come to know something new by remembering it. A salient kind of example involves conceptual enrichment.

Suppose that a young child attends a wedding service with her non-theist parents. Suppose also that, alhough she is fascinated by the Roman architecture of church and the highly structured ceremony, the child entirely lacks the concept **CATHOLIC.** Once home, the child's aunt asks her what she did that day and the child reports, "I went to a wedding." In such a case, it is true that she remembers that she went to a wedding. It is not true, however, that she remembers that she went to a Catholic wedding for she cannot conceptualize her experience in the ways needed to remember *that* fact.

Years later, suppose that the child, now a young adult, is asked whether she has ever been to a Catholic wedding. Having developed competence with the relevant concept while growing up, she is able to answer affirmatively. In doing so, she remembers that she went to a Catholic wedding. And, contrary to a preservative

theory of remembering, she comes to know something new – that she went to a Catholic wedding.

Transformations

A third kind of challenge to preservative theories comes from the fact that memory often transforms information and/or representations between the initial tokening of a representation and the subsequent remembering event. I will describe three familiar ways that this can happen.

Case 5: Condensed memories

Bernecker (2008) notes that many of our memories do not refer to particular events but rather to general periods and events in our lives. Moreover, in many cases we remember such things without being able to remember any particular occasions on which, e.g. one felt a particular way or tokened a particular belief. He offers the following case to illustrate:

> I remember having disliked eating vegetables as a child. Though it is true that I used to dislike vegetables, at no time in my childhood did I (dispositionally) believe that I didn't like vegetables. Rather than entertaining this general idea, what I thought was that I don't like Brussels sprouts, that mangel is too bitter, carrots too sweet, that spinach has a disgusting texture, etc.[9]

This kind of remembering is problematic for a preservative theory because it is one in which memory has collected and transformed information from a variety of token experiences, judgments, and actions. This particular remembering state has no single counterpart in the past and so cannot have merely been preserved from the past. Rather, it is an amalgamation of a whole host of past mental state contents; memory has simply condensed the information from them into a more general judgment. So when Bernecker remembers having disliked eating vegetables, he

[9] Bernecker (2008:148-9). See also, Barsalou (1988) and Price (1952: 351) for a related discussion.

comes to know something he never before entertained. His memory has generated novel knowledge.

Case 6: Tense

A more mundane challenge to preservative theories of memory involves tense. Consider the following everyday scenario. I drive to the airport and park my car in Lot C. I note this and thereby come to believe that my car is parked in Lot C at time t_1. According to a preservative theory of remembering, my memory does its job only if it preserves that belief and makes it available for later use.

Of course, when it is time to go home, my cognitive system will have transformed my prior belief. It can do this in at least two ways. First, it could transform the tense of the attitudinal state such that I remember that my car *was* parked in Lot C at t_1. Transforming belief tense in this way causes problems for some forms of preservative theory.[10] However, a second kind of transformation provides broader challenges to preservative theories. Generally, when I return from trips I remember *where my car is parked*. That is, I remember that my car *is* parked in, e.g. Lot C, at t_2. Memory has not shifted a present tense verb to a past tense verb in this case; rather, it has updated the time to which my relevant belief state is indexed. This is not something that was ever before known, or represented, by me and is thereby in tension with the idea that memory does its job simply by preserving, e.g. prior beliefs.[11]

[10] See Bernecker, 2008: 147 in discussion of "Library" or "Xerox" theories of remembering.

[11] A related, but more general concern is that preservative theories may be in tension with an evolved purpose for memory – namely, its role in enabling subjects to predict present and future states of affairs and events. See De Brigard, 2014; Schacter and Addis, 2007; Suddendorf and Corballis, 2007; and Dennett and Westbury, 2000 for some discussion.

Case 7: Perspective

One final challenge to preservative theories involves a phenomenon already mentioned in chapter 1 – namely, the role of perspective in memory. Specifically, preservative theories of remembering are in tension with cases in which subjects remember an event from the "observer" perspective – one in which they take an external perspective on the remembered events such that their recollection involves an image of themselves as a participant.[12] The problem for a preservative theory is that, in such cases, memory simply does not preserve the perspective one had during the prior experience. Rather, it transforms the representation of the scene so as to represent the same event in a different way – one that involves something that was not included in the original experience.

In sum, the seven cases discussed in this section collectively support the view that memory does not (always) operate ideally by simply preserving information that was previously acquired (and its epistemic status). Often, it operates ideally when it: (i) supports changes in one's relationship to evidence, (ii) adds information to a prior representation, or (iii) otherwise transforms a prior representation in some non-trivial way. The upshot is that memory is not merely, or even primarily, a preservative faculty.

2.1.2 Preservative theories and the nature of memory processing

Our best understanding of how memory processes function also supports the thesis that memory is not exclusively preservative. Specifically, it is now widely accepted

[12] Some have reported switching perspectives in the course of one remembering episode—at one moment remembering an event from the inside and, at the next, remembering it from an external position. See e.g. Hasselmo (2012), Matthen (2010), Sutton (2010), Conway (2009), Robinson and Swanson (1993), Peacocke (1985), Nigro and Neisser (1983), Williams (1973) for discussion.

that memory processes are fundamentally constructive.[13] According to one popular model (cf. Schacter et al., 1998) constituent features of memory representation are distributed widely across different parts of the brain – no single location contains a complete record of any specific experience. Consequently, remembering one's prior experience involves pattern completion. A cue re-activates a subset of the features that comprised the experience. Activation then spreads to the rest of the constituent features of the experience, eventually terminating in the remembering event.

This is not straightforwardly a problem for a preservative view. An advocate of such a view could simply maintain that the distribution and reactivation of these feature patterns is the means by which information is preserved. However, this process is not immune to intervening influences. During retrieval, the reactivated features patterns and subsequent spread of pattern activation is influenced by other general knowledge one has acquired and/or associated with those features in the interim. In short, modifications of memory representations (whether additions, or other kinds of transformations) are not necessarily or even usually pathological. They are the result of the way memory processes operate; they should be expected and welcomed.[14]

As we saw, a variety of everyday cases of remembering involve the generation (or at least transformation) of epistemic features, this is consonant with our understanding of how memory processes work. Moreover it is all in tension with the spirit of preservative theories of remembering. Consequently, we have

[13] See e.g. Bartlett (1932), Roediger (1980), Engel (1996), Schacter (1995, 2007), Bechtel (2008), among many others.
[14] See e.g. Surprenant and Neath (2009), Schacter and Addis (2007) and Engel (1996) for discussion.

good reason to adopt a generative theory instead. In section 2.2, I argue that non-epistemic generative theories of remembering have deep problems of their own.

2.2 Generative Non-Epistemic Theories - Introduction

Non-epistemic generative theories are generally developed in terms of causal relations, i.e. they are causal theories of remembering. While it is natural to think of remembering as the final state of a causal process, and of memory as some sort of set of causal mechanisms (*cf*. Benjamin, 1956: 323), a causal *theory* of memory includes the much stronger claim that causal relations are the *essence* of remembering. Causal theories *define* remembering in terms of the distinctive causal relations token mental states bear to one another.[15]

In its simplest form, a causal theory of remembering holds that a subject remembers something just in case her current mental state is in some (to be specified) way due to or, because of, some past mental state (Bernecker 2008: 17, 2010: 109; Martin and Deutscher, 1966). To keep the discussion manageable, I will focus on causal theories of fact remembering, though most of the criticism raised in the remainder of this chapter could be applied more broadly. A causal theory of fact remembering can be roughly formulated as follows:

(CT) *S* remembers that *p* at t_2 if and only if:

i. *p* is true.

ii. *S* represents that *p* at t_2

[15] Although contemporary causal theories define remembering in terms of causal relations between mental states, they do not generally aim to be giving a conceptual analysis of the meaning of 'memory' or 'to remember'. Much of the historical criticism of causal theories has focused on this latter project (e.g. von Leyden (1961) and Malcolm (1963)) and, to that extent, is orthogonal to the discussion in this chapter.

iii. *S's* representation at t_2 is *suitably causally* connected to a prior representation that p^* at t_1.

iv. p bears an appropriate logical relation to p^*.

Clauses (i) – (iv) are intended to be individually necessary and jointly sufficient for a subject to remember that p. Clause (i) places a factivity constraint on remembering. One cannot remember that p, unless p. Clause (ii) specifies that a necessary condition on remembering that p is that one represents that p in some or other way. Both theses would be shared by almost any other theory of fact remembering.[16]

The third and fourth clauses are the heart of the causal theory. Cause (iii) is a generic formulation of a causal requirement on remembering, and a causal theorist must develop the notion of a "suitable causal connection" in a theoretically viable way. That is, she must explain what makes a causal relation one that is partially constitutive of remembering. Clause (iv) adds a further requirement on the relationship between two token representational states. The representational states at issue in fact remembering have propositional content and, since causal relations say little about the nature of such content, causal theorists have (generally) recognized the need for some kind of "authenticity" requirement.[17] That is, causal theorists generally aim to identify what kind(s) of logical relation(s) hold between p

[16] The nature of a memory representation is itself a difficult issue for causal theorists. While one might be tempted to think that a subject's representation at t_2 should take the form of a *belief* that p, causal theorists have denied this for various reasons (*cf*. Bernecker (2010), Michaelian (2011), and Martin and Deutscher (1966).

[17] The term "Authenticity" comes from Bernecker (2010), though the need to say *something* about relationship between the contents of two token representational states is widely recognized. Michaelian (2011: 333) is a notable exception. He denies that anything substantive can be said about the constraints on memory contents. Of course, to do so is simply to give up the project of giving an informative theory of remembering and one should strive to do better.

and *p** such that two appropriately causally related mental states with those respective contents will yield an instance of remembering.

In the remainder of this chapter, I argue for three claims. First, causal theories are not better motivated than epistemic theories. Second, the most worked out formulation of the causal theory of fact remembering fails to provide a set of sufficient conditions for remembering. Third, it also fails to provide a set of individually necessary conditions. We do not have a theoretically viable characterization of a suitable causal connection, nor do we have an adequate characterization the logical relations that enable remembering. The upshot is that *a causal theory alone cannot supply a set of necessary and sufficient conditions, even if causal relations are necessary for remembering.* Moreover, while it is unclear how a causal theorist could revise the theory adequately, on her own terms, there is no such problem for an epistemic theorist. Consequently, when taken together with the conclusions reached in 2.1, this shows that the only promising theory of remembering is a generative epistemic one.

2.2.1 Motivation

Four central considerations have been offered as support for non-epistemic causal theories of remembering: (i) causal theories distinguish remembering from other mental state types; (ii) causal theories explain the truth of certain counterfactual claims; (iii) causal theories are supported by the empirical sciences; (iv) causal theories are the only viable kind of theory. I argue that all four fail to support a non-epistemic causal theory over an epistemic alternative.

Distinguishing Remembering

A variety of states and processes involve multiple mental states and events that occur at separate times (e.g. inference, imagination, learning, re-learning, testimony). An adequate theory of remembering distinguishes it from other kinds of mental states and processes. Some have argued that to do so, we must accept that a specific kind of causal relation (namely, a contiguous sequence of memory traces) is necessary for remembering. Two kinds of cases have been used to illustrate this point.

The Hypnotist Case[18]

Suppose that, as a young child, Craig was briefly separated from his family during a trip to Disneyland. The experience was traumatic and for quite some time Craig remembered that he got lost at Disneyland as a child. However, suppose that as a teen, Craig suffered an unusual brain injury that entirely wiped his memory of the trip to Disneyland and of the fact he had been lost there. By chance, however, he took part in a hypnotist's entertainment performance as an adult and had the belief that he was lost at Disneyland as a child implanted by the unknowing performer. Now suppose that when later asked about his life, Craig reports that he remembers that he got lost at Disneyland when he was a child.

Since Craig has this true belief about his past by pure coincidence (the hypnotic suggestion bears absolutely no relevant relation to any events in Craig's childhood) this is not plausibly a case in which Craig genuinely *remembers* that he was lost at Disneyland as a child. An advocate of causal theories attributes the fault in this case to the fact that there is no relevant causal connection between Craig's

[18] This case is derived from Martin and Deutscher, 1966:174.

present belief and his relevant prior beliefs and experiences. Consequently, she takes this to show that a causal condition is needed to distinguish genuine cases of remembering that *p*, from cases in which one comes to have a true belief about the past in some other way.

The Testimony Case[19]

Suppose again that Craig visited Disneyland and was briefly separated from his family. Now, suppose that in addition to remembering that he got lost at Disneyland, Craig shared this fact with his friend Peter. Again, an unusual brain injury during his teens wiped his memory of the fact that he had been lost. Later, suppose that Peter told Craig that he had been lost at Disneyland as a child, and Craig came again to believe as much. Finally, suppose that Craig forgot that Peter told him about this fact, but nevertheless retained the belief.

In this case, Craig's belief about the past is not true by chance. There is a causal connection between his prior experiences and corresponding mental states, and his present belief. However, since Craig's brain injury led to his *forgetting* the event and corresponding fact, the causal connection goes by way of a testimonial chain involving Peter. Craig would not have formed a true belief about his past at t_2, without having learned of the fact about his past from Peter. Given that it involves multiple people and multiple minds, a causal theorist might conclude that this is the *wrong kind* of causal connection between Craig's present belief and his past experiences. Consequently, she might be tempted to conclude that we need a *specific*

[19] Martin and Deutscher, 1966: 174 use a similar case for a slightly different end.

kind of causal condition to rule out cases of relearning (e.g. one that restricts the causal path between mental states to internal and contiguous memory traces).[20]

The cases do not support a pursuit of non-epistemic causal theory of remembering over epistemic alternatives. First, since any adequate theory of fact remembering should distinguish it from other mental state types, this need equally motivates a generative epistemic account of remembering. Moreover, nothing in an epistemic account of remembering requires that there be *no* causal connection between prior experience and present mental state. The central question is whether requiring a *particular* causal requirement is the best way to distinguish remembering from other mental states. Careful examination of the cases described above does not support that conclusion.

Consider the hypnotist case. According to the causal theorist, we need to posit a causal condition to prevent cases in which one's true belief about the past bears no causal connection to the past events/facts of which they are about from qualifying as instances of remembering. This is plausible, however, one can ask why this might be. What is problematic about a case in which fortuitous circumstances lead one to luckily hold a true belief about her own past, and why is a causal link the crucially absent element in such cases?

In isolation, these questions are somewhat difficult to answer. However, in the context of an epistemic theory there is an explanation readily available. The problem is not merely that there is no relevant causal connection between the event and a subject's true belief; rather, the problem is that *because* there is no relevant

[20] See, e.g. Bernecker, 2010: 106.

causal connection, one's true beliefs are only coincidentally true and such beliefs cannot be instances of knowledge. Since, certain versions of epistemic theory require that remembering that *p* entails knowing that *p*, they explain why such cases cannot be instances of remembering. Establishing a causal connection between prior events/experiences and present beliefs is important *only* because it is a way for a subject's true belief about the event to be epistemically virtuous. Consequently, the epistemic theorist has a principled explanation of why hypnotist-style cases fail to be instances of remembering, and has an explanation for why we should require some or other (but perhaps no unique kind of) causal connection between prior experiences and present true beliefs about those experiences. The upshot is that such cases provide at least as much (if not more) motivation for an epistemic theory of remembering than they do for a causal theory of remembering.

The testimony case makes this even clearer. The worry, for someone drawn to a non-epistemic causal theory of remembering, is that this must be an instance of re-learning and not remembering because the causal connection between prior experience and present belief is of the wrong kind.[21] However, it is implausible that *no* instances of fact remembering can involve a testimonial chain. In fact, much of what we remember, whether general facts about the world or particular facts about our own lives, seems to depend crucially on testimonial chains.[22] For example, I remember that my own birth was complicated and that the attending physician used forceps to pull me out by the ears. Of course, I can only remember these facts because my parents shared them with me much later in my life.

[21] See e.g. Bernecker (2010: 129).
[22] See Harris et al. (2013), Michaelian and Sutton (2013), and Sutton, (2014, 2011, 2008).

So it seems that the causal theorist faces a dilemma. On the one hand, she can hold that fact remembering cannot involve intervening testimonial chains. This has the unacceptable consequence that many of the facts we seem to remember are not things we remember at all. On the other hand, she can allow that fact remembering can go by way of testimonial chains as well as by causal chains that are internal to the subject. This has two undesirable consequences (from the causal theorist's perspective). First, it follows from this that there is not a *unique* type of causal connection that is distinctive of remembering. If there is no single type of causal chain that is partially constitutive of remembering, then the causal theorist needs a non-ad hoc explanation of the shared feature of these various causal chain types that makes them of the right kind for remembering. Second, she must explain what, if anything, about the original testimony case provides support for a causal theory.

To respond to the second challenge, the causal theorist might start by drawing a distinction between two different things Craig might be said to remember. In the original testimony case, the question was whether it was plausible that Craig remembered *that* he had gotten lost at Disneyland as a child. The example in which testimony is crucial to my remembering a particular fact about my own birth showed that there really is no challenge to the idea that Craig might be said to remember that he had gotten lost on the basis of Peter's testimony. However, the causal theorist might maintain that there remains something that Craig cannot remember in these circumstances; namely, he cannot remember *getting* lost as a child. The question now is whether, having drawn this distinction between remembering *that* something happened, and remembering it *happening*,

provides motivation for a causal condition on *event* remembering in ways that it does not also support an epistemic theory of remembering.[23]

It is not clear that it does. In fact, the epistemic theory appears to provide an even better explanation of the cases in light of this distinction. According to an epistemic theory, it should not be surprising that Craig can naturally be said to remember that he had gotten lost in this case. There is nothing to suggest that Peter's testimony cannot sustain the transmission of knowledge across time. Consequently, it is natural to say that Craig knows that he was lost as a child on the basis of his memory, which has Peter's testimony, and not merely the original experience, as its source. So, the causal and epistemic theories are on equal footing with respect to a particular fact that Craig remembers about his own life.

The causal theory denies that Craig can remember getting lost because the causal connection between his present mental state and the event is of the wrong type. Explaining *why* the causal chain is of the wrong type is, once again, difficult in isolation. The causal theorist needs a way to distinguish between deviant and non-deviant causal chains and this is a notoriously difficult problem. However, notice that this problem does not (directly) emerge if one endorses an epistemic theory of remembering for there is a relatively straightforward diagnosis of this case. One kind of epistemic theory (the one I advocate in chapter 3) delivers the verdict that

[23] Bernecker (2010:131) essentially acknowledges that fact remembering can be transmitted via testimonial chains in the way described here. However, he draws a distinction between remembering what a testifier said about an event and the event itself, and denies that one can remember the event on causal grounds. This is problematic for two reasons. First, that someone's testimony about an event is the source of one's remembering does not guarantee that one remembers what the testifier said about the event. Indeed, one could conceivably be (and perhaps we often are) wrong about what a testifier literally *said* while still getting the facts right. Second, as we saw in chapter 1, it is possible to remember objects and events without a causal chain tracing back to one's experience of those objects and events.

Craig cannot remember *getting* lost because remembering is a matter of knowing by operations of one's memory and the kind of intimate first-personal knowledge one has in remembering one's own experience of an event is not the kind of knowledge one can acquire from a second or third person testimonial report.[24] The upshot is that an epistemic theory of remembering is equally, if not better, motivated by the need to distinguish remembering from other mental state types, and by the cases used to illuminate this need.

Counterfactual Explanation[25]

A second motivation for the causal theory of remembering is derived from the fact that we often explain the occurrence of memory states in terms of counterfactuals of the form, 'If *S* hadn't represented that *p** at t_1, he wouldn't represent that *p* at t_2'. For example, suppose that Craig genuinely remembers that he was lost at Disneyland as a child. We think that he wouldn't believe that had been lost at Disneyland at t_2, if he had not believed (or at least represented) that he was lost at Disneyland at t_1 (perhaps while being lost). Now consider the following two claims:

(i) Craig would not believe that he had been lost (at t_2) unless he had believed that he was lost (at t_1).

(ii) Craig's having believed that he was lost (at t_1) is the cause of his believing that he had been lost (at t_2).

[24] Note that this does not commit us to the view that one cannot, in principle, acquire first-personal knowledge of the past by way of *some* external source – more anon in discussion of Bernecker's causal theory in section 2.2.3.

[25] See Bernecker, 2010:104, 124-7 for this particular argument. See Martin and Deutscher, 1966: 176 for a related discussion.

Bernecker (2010: 125-6) maintains that causal statements like (ii) provide the best explanation of the truth of counterfactual statements like (i). He reasons that a *non-causal* account of counterfactuals like (i) will fail to explain the *process* that underlies the retention of knowledge or representations. In contrast, a causal account does at least gesture at an *explanation* of *how* it is possible that, e.g. Craig's ability to represent that he was lost as a child is retained from t_1 to t_2.

Let us grant Bernecker that such counterfactuals are best explained causally. This demonstrates only that remembering that *p* at least partially causally depends on prior events and representations. It offers no support for any *particular* causal relation. Moreover, there is no reason to deny that one's memory operations involve some or other causal processes. Consequently, such considerations also support a generative epistemic theory of remembering that holds that, e.g. one remembers that *p* at t_2 just in case she knows that *p* at t_2 by operations of her memory. Further, according to an epistemic theory, the observed causal dependencies can be said to emerge from a familiar conception of knowledge.[26] In contrast, for the causal theorist, positing such dependencies is merely an attempt to do justice to particular cases. On balance, this appears to favor an epistemic theory of remembering.

Empirical Science

A third consideration offered in favor of a causal theory is that it is, supposedly, the theory that is best situated to make sense of the picture of memory processing that is emerging from the empirical sciences. In particular, some claim

[26] This need not be restricted to causal theories of knowledge of the form developed by Goldman, 1967. The observed causal dependencies would emerge from a variety of theories of knowledge that allow for causation to play an important role.

that it is best suited to make sense of the *constructive* nature of memory. For example, Michaelian (2011: 325) argues that a causal theory of memory best makes sense of the following features of memory processing (*cf.* Alba and Hasher (1983):

(i) Selection: Only certain incoming stimuli are selected for encoding.

(ii) Abstraction: The meaning of a message is abstracted from the syntactic and lexical features of the message.

(iii) Interpretation: Relevant prior knowledge is invoked.

(iv) Integration: A holistic representation is formed from the products of selection, abstraction, and interpretation processes.

(v) Reconstruction: During retrieval, whatever information was selected for representation and is still accessible is used, together with general knowledge, (roughly) to generate a representation of must have happened.

These principles can be illustrated with the following case. Suppose that a subject witnesses a car accident at t_1. Although her visual experience is tremendously rich, only some of the visual stimuli are selected (e.g. the shapes and colors of the two involved vehicles, and so forth) by her cognitive systems for further processing. The selected information is abstracted and interpreted using available prior knowledge and all of this is integrated into a holistic representation of the event (e.g. she comes to have a representation of an accident event involving one blue and one red car). Suppose that the subject is later (at t_2) shown images of the accident wreckage. Her cognitive systems select, abstract, and interpret the information available in that experience to form an additional representation of the

seen images. Now suppose that she is asked the following (at t_3): "What directions were the respective vehicles heading at the time of the accident? Which vehicle initiated the impact? And what was the point of impact?"

In an effort to answer these questions, the subject tries to remember the accident. Her memory processes piece the information that had previously been encoded and interpreted at times t_1 and t_2 (or as much of it as remains available) together with her other general knowledge about physical interactions (e.g. she knows that an object that has a dent on the left front fender has been impacted from the (relative) left front direction, etc.) to *construct* a representation of what happened. As a result, she comes to remember *that* the blue car was headed north; the red car was headed south; and the red car crossed the center lane impacting the blue car at the left front fender.

According to a causal theorist, the best way to make sense of all of these features of memory processing is to take causal relations to be the essence of remembering. Since memory processes modify information in a variety of ways, the one constant, one might think, is the causal relation between prior and present representations.

However, such a conclusion would be premature. Constructive memory processes are not in tension with the generative epistemic theory of remembering. An epistemic theorist simply understands constructive memory processes to be the mechanisms by which one comes to know that *p* by using one's memory. In particular, one should not be led astray by the claim that memory generates a "representation of what must have happened." In the same way that not all outputs

of a memory system are created equal (some are genuine instances of remembering and others are at best instances of *apparent* remembering) not all hypotheses are created equal either. Any adequate theory of remembering must provide a means for distinguishing between good and bad memory outputs, i.e. it must distinguish genuine remembering from merely apparent remembering. I will argue below that the causal theory does not have adequate resources to do this in any satisfactory way. In contrast, a generative epistemic theory provides a principled framework—selection, abstraction, interpretation, and integration processes contribute to genuine remembering only when their outputs are instances of knowledge. The upshot is that a generative epistemic theory of remembering is at least as well supported by the constructive nature of memory processes as any causal theory would be.

An Indirect Argument

Finally, Bernecker (2010: 104-114, 2008: 18) maintains that one of the most compelling motivations for the causal theory of remembering is that it is simply less problematic than all available alternatives. Of course, every view is motivated by the conviction that it outperforms all of its competitors and so this consideration motivates the generative epistemic account as well. It is, however, dialectically useful to see why Bernecker takes the causal theory to be the least problematic kind of theory.

Bernecker envisions three characterizations of the memory connection between representations at t_1 and t_2: a 'simple' connection, an epistemic connection, and a causal connection. He argues that only the causal condition is plausible.

He formulates and assesses a '"simple" theory as follows:

(SR) *S* remembers that *p* at t_2 just in case (i) *S* represents that *p* at t_2, (ii) *S* represented that *p* at t_1, (iii) and *S* has not forgotten that *p* in the interval between t_1 and t_2.

This simple proposal is intended to be an alternative to a causal account of remembering. That is to say, it is a theory in which remembering is not distinguished from other mental states via appeal to any causal condition.[27] The idea is that remembering is distinct from other mental states in virtue of being a representation that one had previously acquired and that one has not since forgotten. Setting aside the fact that this proposal is *prima facie* committed to a preservative theory of remembering, Bernecker offers two further objections.

First, it is difficult to see what exactly is meant by "forget" in this context. We often seem to at least temporarily forget something that we later remember only in virtue of receiving the right kinds of prompts (Bernecker, 2010: 110-11). Of course, one may be able to finesse the proposal to accommodate this observation. Nevertheless, Bernecker thinks there is a second, much deeper problem for this simple proposal. Namely, it does not have the resources to answer some of the most pressing questions regarding the memory relation. The simple theory tells us only that one has retained some ability to deploy a representation that *p*; it says nothing about what is involved in the process of retaining such an ability. Since Bernecker doubts that the idea of non-causal information storage even makes sense (112), he thinks any attempt to develop the simple theory will collapse into a causal theory of

[27] Squires, 1969; Ryle, 1949; Munsat 1967; and arguably Zemach, 1968 endorse versions of this thesis.

remembering. Thus, the causal theory is less problematic than this simple theory of remembering.

I share Bernecker's concern that such a theory lacks explanatory power. However, the same objection would not apply to an epistemic theory because the epistemic theory can allow for, or even presuppose, some causal explanation. Moreover it offloads some of the explanatory demand to theories of knowledge. So, rejection of a simple connection principle supports non-epistemic causal theories and epistemic theories of remembering equally.

This leaves an epistemic theory and a causal theory as available options. Bernecker provides the following formulation and assessment of an epistemic theory of remembering:

A subject *S* remembers that *p* iff

i. *S* knows that *p* at the time of remembering (time *t*).

ii. *S* knew that *p* at some prior time (time *t'*).

iii. *S's* knowing that *p* at *t* is preserved from *S's* knowing it at *t'*.

Bernecker rejects conditions (i) through (iii) as necessary conditions for remembering that *p* (2010: 65-103, and 105-109). Of course, a generative epistemic theory of remembering is not committed to conditions (ii) and (iii). Consequently, the causal theory can only be shown to be less problematic than *all* of its epistemic competitors, if one can show that one can remember that *p* without knowing that *p*. I defer discussion of this issue to chapter 5 (section 5.4). That qualification aside, the upshot is that it is not at all clear that a non-epistemic theory of remembering is better motivated than an epistemic theory of remembering. Moreover, I turn next

the causal theory itself and I argue that it has deep problems of its own. This suggests that, if anything, a generative epistemic theory of remembering appears to be the least problematic kind of theory of remembering.

2.3 The Causal Theory of Remembering

An adequate causal theory of remembering provides a set of necessary and sufficient conditions for remembering. In the course of doing so, it does two key things. First, it characterizes the kind of causal relation that is partially constitutive of remembering. Second, it identifies the minimal logical relation or relations that must hold between a candidate remembering state and one's prior mental states such that two appropriately causally related mental states with those respective contents will yield an instance of remembering. The most thoroughly developed causal account of fact remembering can be found in Bernecker's (2010) where he offers the following:

S remembers that *p* at t_2 only if:

i. *S* represents that *p* at t_2

ii. *S* represented that *p** at t_1

iii. *P* is true at t_2

iv. *P* and *p** supervene on the same environmental conditions at t_1 or *p* is entailed by *p** (where entailment is understood along the lines of relevance logic).

v. *S's* representation that *p* at t_2 is causally connected to *S's* representation that *p** at t_1 such that

a. *S's* representation that p^* at t_1 and *S's* representation at t_2 that p are connected by a persisting memory trace or a contiguous series of memory traces.

b. The memory trace is at least an inus condition for *S's* representation that p at t_2. If the memory trace is an independently sufficient condition, it is not preempted by another independently sufficient condition.

c. If *S* hadn't represented that p^* at t_1 he wouldn't represent that p at t_2.

On Sufficiency

Bernecker formulates this account as a set of necessary conditions for remembering. And one can appreciate the desire to remain modest. Many philosophical analyses have been subject to counterexamples directed at the sufficiency of their analysans, and Bernecker is sensitive to this possibility for his own proposal.[28] However, it is clear that he intends this set of necessary conditions to provide as close to an account of fact remembering as he thinks is possible. Unfortunately, Bernecker's proposal does not provide a set of sufficient conditions and it is instructive to see exactly what its shortcomings are. Consider the following two cases.

Case 1: Unjustified Memory Beliefs

Suppose that one morning (t_1) I simply wake up with the belief that there is life on Planet G. Never have I received any evidence in favor of this claim, and I don't particularly care whether or not there is life elsewhere. I believe it for the sake of

[28] Personal Correspondence.

believing something. Suppose that there is life on Planet G and, at t_2, I find myself with the occurrent belief that there is life on Planet G. Now suppose that my belief at t_2 depends on my belief at t_1 in all the ways normally associated with remembering. In such a case, my belief at t_2 meets conditions i-v. However, it should be clear that I do not actually remember that there is life on Planet G. The world itself is the domain of fact remembering and so such remembering must connect to the world. We cannot willfully (or unintentionally) remember that *p* just because we wanted to believe that *p* at some point, and then had our memory processes function "normally". What we can remember about reality must be constrained by reality, and Bernecker's account does too little to secure an appropriate relation between the world and a subject's representations of it.

Case 2: Lucky Memory Beliefs

Russell's Clock

> Suppose that one morning Jo comes in to class and looks up at the clock to check the time. It reads exactly 8:22, and as the clock had never led Jo astray before, she is justified in forming beliefs about the current time on the basis of its testimony. She does and comes to believe truly that it is 8:22. Suppose further that she gets the time right by mere luck for the clock had stopped working at exactly 8:22 the previous night. Later that night, Jo tries to recall when she made it to class and on the basis of her prior belief and functioning memory processes forms the justified true belief that she came in at 8:22.

This case shows that the problem is not merely a matter of forming unjustified beliefs. Jo's beliefs meet conditions i-v, but here too it should be clear that this fails to be an instance of fact remembering. Though she has reason to believe that she came in at 8:22, Jo's memory belief still bears no relevant relationship to the fact that makes it true and no subsequent amount of memory

processing can create a relevant relation out of nothing. The analysis will simply not do.

In response, a causal theorist might be tempted to strengthen the relationship between memory and the world by adding a two conditions to the analysis: (i) a 'reliability condition' that specifies that the cognitive processes that support a memory belief must be reliable and (ii) a 'content source condition' that constrains the possible sources of the content of one's memory belief such that one only remembers if one's original belief was formed on a suitable basis, e.g. perception or inference.[29]

There are at least two problems with these additions. First, both appear to be ad hoc revisions. If the essence of remembering consists in causal relations between mental states at times, we need a principled reason to go beyond the commitment that the causal relations involve memory traces. We also need a principled reason to provide any restriction on the possible sources of memory contents. Compare the state of being hit by a ball. Such a state can be adequately characterized in causal terms – being hit by a ball is only a matter of certain causal relations between that person and a ball. That is, one can provide a set of sufficient conditions for being hit by a ball by specifying a set of causal conditions between ball and individual. Adding the analogous reliability and source constraints to such an account would be entirely unmotivated. It does not matter what the source of the ball's motion is, or whether the causal events leading up to the ball's impact with the subject are in any sense reliable.

[29] Michaelian adopts both conditions in his preferred formulation of a causal theory (2011: 235-6).

Second, the notions of "reliable processes" and "suitable sources of belief" that have been suggested are epistemic in nature. Consequently, adopting such principles simply abandons the spirit of a non-epistemic account of remembering. The cases described above plausibly show that an account of remembering that *p* needs something in the form of additional constraints. The problem is that the candidate supplementary constraints support the conclusion that remembering is primarily a causal matter, but rather, it is fundamentally an epistemic matter. The upshot is that the most developed causal theory of fact remembering fails to provide a set of jointly sufficient conditions for remembering that *p*, and it is difficult to see how one might do so in any principled way on the causal theorists own terms.

On Necessity

The discussion so far would suffice to motivate pursuit of a generative epistemic theory of remembering. However, it is not clear that Bernecker's particular proposal provides a set of necessary conditions for remembering either. Specifically, clauses (iv) and (v) need not be met by some instances of fact remembering.[30]

I begin with the causal condition (clause (v)). Epistemic and non-epistemic theories can agree that some or other causal processes support memory. The crucial difference is that for an epistemic theory, the kind(s) of causal processes that do so are not at the heart of the analysis. Causal connections are simply enabling or explanatory conditions for something else that is partially constitutive of remembering (e.g. knowledge). In contrast, the formulation of a causal condition is

[30] Bernecker actually acknowledges this point (e.g. 2010: 243) but does not, to my mind, appreciate its importance.

at the heart of a causal theory because a causal theory is distinctive only in so far as it takes a certain kind of causal connection to be constitutive of remembering (Bernecker, 2010: 128). Given that, recall Bernecker's formulation of a causal connection:

S's representation that *p* at t_2 is causally connected to *S's* representation that *p** at t_1 such that

i. *S's* representation that *p** at t_1 and *S's* representation at t_2 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 condition for *S's* representation that *p* at t_2. If the memory trace is an independently sufficient condition, it is not preempted by another independently sufficient condition.[31]

iii. If *S* hadn't represented that *p** at t_1 he wouldn't represent that *p* at t_2.

This merits some explanation. The clause has three distinct components, each of which must be satisfied if one is to remember that *p*. (i) The causal relation must go by way of a persisting memory trace or contiguous series of memory traces (engrams). (ii) Those traces must not be redundant components in the production of the representation at t_2. (iii) The resulting causal connection must support certain counterfactuals.

There are at least two problems with this proposal. First, as at least one causal theorist has already recognized, this kind of causal condition is in tension

[31] An inus condition for x is an insufficient, but non-redundant factor of an unnecessary but sufficient condition for x cf. Mackie (1965).

with the contemporary science of memory.[32] Remembering involves ubiquitous content transformation and content incorporation from a variety of resources at a variety of stages (e.g. background knowledge, intervening testimony, and experience, etc.). Consequently, it is at least (in principle) possible that there is no single memory trace, or series of memory traces, that meets all three of these conditions despite the fact that one remembers that *p*.

For example, we could envision a version of condensed memory case in which Bernecker's remembering 'that he disliked vegetables as a child' is the causal product of a collection of *redundant* traces and cues, some containing information about his reaction to Brussels sprouts; others his reaction to peas and carrots; and so forth. In such a case, each individual memory trace would arguably fail to be an inus condition for his present representation because each would be causally redundant. Moreover, it is not clear that any of these memory trace connections would support the counterfactual Bernecker requires. Consider the trace associated with his prior judgment that "carrots are too sweet". Given the collective role of the other traces, it is false that, if Bernecker hadn't represented that carrots are too sweet at t_1, he wouldn't represent that he disliked vegetables as a child at t_2. The upshot is that remembering that *p* need not satisfy the causal condition as formulated.

[32] Michaelian 2013: 2444. Despite acknowledging this problem, Michaelian appears to be at a loss for what to do about it. He considers the possibility that one could modify the causal theory such that it does not require a continuous causal connection via distributed memory trace but, to my mind, he prematurely rejects the possibility on the grounds that it might commit one to an epistemic theory of remembering (2013: 2452).

In response to this objection, one might be tempted to weaken the causal condition such that no individual memory trace must be an inus condition or support the relevant counterfactual. Of course, doing so threatens to diminish the impact of the proposal as a non-epistemic account of what it is to remember that *p*. The less a theory commits to with respect to the nature of the causal relations between remembering and prior mental states, the more plausible it is that causal relations really aren't at the heart of the matter.

A second problem with this proposal is that, though they typically underlie remembering, memory traces are not plausibly *essential* to the remembering process. To see why, we have to get a bit clearer on exactly what a memory trace is. A standard definition is as follows:

(MT) The neural changes that accompany a mental experience at one time (t_1) and whose retention, modified or otherwise, allows the individual later (t_2) to have mental experiences of the kind that would not have been possible in the absence of the trace. (Tulving, 2007b: 66)

Tulving is likely operating with a broader conception of "mental experience" than is familiar in philosophy. In particular, it is not clear that the relevant mental states at t_1 and t_2 need be conscious in the way philosophers typically characterize experiences as being. This qualification aside, the definition highlights three key features of memory traces. First, they are changes in neural networks. Second, they are distributed across many brain regions (Moscovitch, 2007: 21). Third, they are dynamic, changing, malleable entities (Dudai, 2002; Nadel, 2003). That is, they are crucially open to modification by future mental (and neural) events, and their being

so allows the organism to incorporate new experiences into existing knowledge structures (Nadel, 2007: 180).

Given this characterization of memory traces, it is clear that they are not, strictly speaking, necessary for remembering that *p*. The gist of the problem is that any proposal that makes essential use of memory traces will be inadequately flexible. For example, according to this proposal it is metaphysically impossible for something to remember in the absence of it having neural networks. Second, and much worse, it makes it metaphysically impossible for humans to remember in ways differently than they typically do. Suppose that neuroscientists developed the ability replace neural structures as they began to fail with functionally equivalent non-biological components (e.g. silicon chips). Suppose that they got so good at doing so that they could replace dying or malfunctioning neural components with non-biological components that functioned in exactly the same ways as healthy neurons did. As formulated, this proposal has the consequence that, despite the fact that such replacements would, in principle, enable subjects acquire and retain new information in ways that perfectly parallel their prior abilities, post-surgery subjects would be entirely incapable of remembering anything. After all, the causal mechanisms that underlie the persistence of information in such subjects do not involve memory traces at all.

Similarly, suppose that scientists discover how to seemingly eliminate some of the amnesia often associated with various medical procedures by transferring information from one's neural networks into digital networks and then "re-uploading" it afterward. Regardless of the behavioral results, this proposal would

entail that the scientists were completely wrong. They were unable to eliminate any amnesia. In fact, doing so would result in the complete loss of all memories for the subject, as the digital network would neither be composed of neural networks nor likely store information in the same spatially distributed fashion.

The point of these cases is not that there are any creatures that have the capacity to remember despite lacking a neurological system. Nor is it that scientists will actually succeed in functionally duplicating one's neural structures or "downloading" the information encoded in one's neural networks. Rather, the point is that these should at least be metaphysical possibilities, and the causal theory, as formulated, does not allow for this.[33]

The causal theorist can respond to this objection in two ways. First, she can simply accept the consequences of her view and try to re-describe the above case as involving something much like remembering but not remembering itself. On the one hand, doing so allows her to retain one of the most important virtues of the causal theory – that it gives a characterization of the kind of causal connection that is partially constitutive of remembering. On the other hand, the move appears to be ad hoc and done for the express purposes of preserving one's favored theory.

Alternatively, the causal theorist can re-characterize the notion of a memory trace in more abstract terms. Bernecker chooses this route (2010: 137-140). He characterizes memory traces as intracerebral occurrences, only in the sense that they are dispositional representations (or otherwise subdoxastic states) that are

[33] It is also worth noting that the current formulation of the proposal makes the notion of "collective memory" a metaphysical impossibility as well. I will not pursue this issue here but this does appear to be another significant cost of the theory.

typically realized in the brain.[34] Consequently, there is no, in principle, reason to deny that the above cases are instances of remembering.

I am sympathetic to this response. However, it is not particularly helpful for a *causal theorist*. The problem is that it effectively acknowledges that there is a plurality of possible causal mechanisms that support remembering. This mitigates the explanatory value of the proposal qua causal theory. A causal theory of remembering aims to describe a causal connection condition that is partially constitutive of remembering, and in doing so, to provide a means by which to identify deviant causal connections. Adopting a traditional conception of memory traces goes some distance toward accomplishing this – the memory connection is the one that involves changes in certain neural structures and has features that continue to be uncovered by contemporary memory sciences.

Reconceiving memory traces in a manner that is consistent with a plurality of causal mechanisms (some distributed across brain regions, some involving information encoded in a digital network, etc.) makes identifying deviant causal chains considerably more difficult. Once one grants that more than one causal mechanism can support remembering, one is pressured to identify what those different mechanisms have in common such that they support remembering, and what other similar relations lack such that they are to be considered deviant. In sum, reconceiving of memory traces in the suggested way suggests that there is something deeper, more essential to remembering, than the causal mechanisms themselves.

[34] Specifically he adopts the non-reductive physicalist framework that emerges from Pinker and Prince's (1988) "implementational connectionist" model.

The significance of this point is made clear by revisiting some example cases. Bernecker's characterization of memory traces is largely motivated by the desire to allow for cases in which information is not stored exclusively in neural networks but, e.g. instead downloaded into a digital network and later re-uploaded back into the brain and so forth (cf. Bernecker 2010: 139-40). Yet, part of the motivation for a causal theory in the first place was to rule out cases in which information is transmitted via chains that are external to a subject's cognitive systems, e.g. to rule out remembering that involves a testimonial chain (cf. Bernecker, 2010: 130). Granting for the sake of argument that one remembers in the memory transplant case and fails to remember testimony case, the causal theorist must go on to explain which difference(s) justify the differential assessment. And here the causal theorist has little to appeal to besides brute intuition.[35]

Similar points apply to condition (iv)). Recall that it was formulated as follows:

> *P* and *p** supervene on the same environmental conditions at t_1 or *p* is entailed by *p** (where entailment is understood along the lines of relevance logic).

Bernecker includes the supervenience requirement to explain why Twin-Earth scenarios can lead one to fail to remember that *p*. Specifically, the idea is that one will fail to remember that *p* when the content of the concepts one deploys in representing that *p* at t_2 supervenes on the social and environmental facts on earth, while the content of the concepts one deployed in representing *p** at t_1 supervened

[35] Bernecker's judgments do seem to be guided by intuitions, though I confess I do not always share them (more anon in chapter 5). See his 2010: 137-40 for some examples.

on the social and environmental facts on twin-earth. Though sympathetic to this aspect of the proposal, I set it aside.

The more interesting issue concerns the entailment relation. Bernecker introduces the entailment condition in an attempt to characterize what it is for a memory content to be similar enough to the content of one's prior mental state so as to count as an instance of remembering (when all the other conditions are met).[36] Bernecker has the right idea here, but the proposed relation is too demanding. One can remember that *p* at t_2 in cases in which *p* is not *entailed* by any prior mental state or states. We have already seen several cases in which this is so.

For example, cases in which remembering involves conceptual enrichment need not satisfy an entailment condition. Suppose a small child believes that she is at a wedding at t_1, acquires the concept of Catholicism at t_2 and later (at t_3) remembers that she went to a Catholic wedding. Her belief at t_1 had the following content: [I am at a wedding]. Her remembering state at t_3 had the following content: [I went to a Catholic wedding]. The former simply does not entail the latter. Concept acquisition need not go by way of entailment.

Similarly points apply to certain cases involving tense. Recall the case in which I remember where my car is parked upon returning from holiday. At t_1, I formed the belief that my car is parked in Lot C (at t_1). At t_2, I remember that my car is parked in Lot C (at t_2). The former does not entail the latter.

In both of these cases there is some kind of logical relation between the contents of the two mental states, but it is not specifically an entailment relation.

[36] Bernecker, 2010: 222

Consequently, Bernecker's proposal does not adequately characterize the conditions under which two mental states have contents that are similar enough so as to be memory-related to one another.[37]

In response to this problem one might weaken the logical relation one takes to be necessary for remembering that *p*. One might even accept that there is a plurality of candidate logical relations that will do. Of course, once one does so, one can ask what each of these logical relations has in common that they can support remembering and, once again, the causal theorist has no ready answer.

2.4 Conclusion

The upshot of this chapter is that we have good reason to pursue a generative epistemic theory of remembering. Preservative theories are subject to plausible counterexample cases and are in tension with the way that human memory processes work. Generative non-epistemic theories (i.e. causal theories) of remembering have deep problems of their own. Though some or other causal processes are likely necessary for remembering, none would be sufficient. Moreover, the best going causal theories fail to adequately characterize the causal and logical relations that are supposed to be partially constitutive of remembering. More resources are needed.

[37] Though Bernecker focused exclusively on fact remembering, it is worth noting that this problem is greatly exacerbated once one attempts to extend the proposal to non-propositional remembering states or remembering that involves a combination of propositional and non-propositional content. For example, on might correctly self-report that, "I vividly recall the first tremor of the earthquake, not that, at the time, I realized that it was an earthquake – I didn't even know what an earthquake was". Or, "I remember the priest officiating the ceremony, not that I knew that he was a priest or what it was to officiate a ceremony at the time."

Chapter 3 A Generative Epistemic Theory of Fact Remembering

In Chapter 1, I argued that there were four basic kinds of theories of remembering. In Chapter 2, I argued that three of those four are not adequate. Preservative theories of remembering, epistemic or otherwise, fail to capture a variety of data and are in tension with an emerging consensus about the nature of actual human memory processing. Causal theories fail to provide a set of sufficient conditions for remembering, and the best formulations fail to identify an adequate set of necessary conditions as well. This leaves a generative epistemic theory as the last available kind of theory of remembering. The goal in this chapter is to describe and motivate a particular version of generative epistemic theory of fact remembering. It takes the following form: *S* remembers that *p* at *t* iff *S* knows that *p* at *t* by operations of her faculty of memory. Characterized in this way, memory is a genuine source of knowledge, sometimes on par with other sources like perception and inference. Its success state – remembering – is a species of knowledge.

In section 3.1, I describe two possible formulations of a generative epistemic theory: a strong theory, as just described, and a weak theory, according to which fact remembering is a species of justified belief. In section 3.2, I describe five arguments for a generative epistemic theory of fact remembering and argue that they best support the stronger of the two theses. In section 3.3, I show that the view successfully deals with key cases discussed in chapter 2. And, in section 3.4, I discuss the notion of "knowing by operations of one's faculty of memory". In the course of doing so, I also respond to a preliminary objection to the theory on the grounds of circularity.

3.1 Formulations of G.E.T. – Strong vs. Weak Theses

The argument by elimination in chapter 2 motivated pursuit of a generative epistemic theory of remembering. Though it supported the conclusion that remembering is fundamentally a kind of epistemic success, it gave no indication of what that success amounts to, i.e. what an epistemic a theory should look like. There are two basic ways to develop an epistemic theory. First, one could characterize remembering as a paradigmatic instance of epistemic success – most naturally, one could take it to be a species of knowledge. Second, one could characterize remembering as some lesser epistemic success. In the case of fact remembering, one could naturally take it to be a species of justified belief. I will label the former view a *strong* epistemic theory and the latter view a *weak* epistemic theory.

A strong generative epistemic theory takes the following generic form:

(GET) A subject *S* remembers something *x* at *t* iff *S* knows *x* at *t* by operations of her faculty of memory.

In this generic form, the thesis is neutral with respect to the objects of remembering and can potentially be adapted to give an account of the remembering of facts, objects, and events. It is a *strong* epistemic thesis because the relation is one in which remembering is naturally taken to be a *species* of knowledge – remembering something is sufficient for knowing it. It is a generative thesis because it includes no requirement that one have previously known *x* to remember *x* at *t*. Consequently, when one remembers, the operations of one's memory processes may generate *novel* knowledge.

A weak epistemic theory of remembering takes the following generic form:

(GET*) A subject *S* remembers something *x* at *t* iff *S* bears some epistemically successful relation to *x* at *t* by operations of her faculty of memory.

This shares many features with the strong thesis. However, it is a *weak* epistemic thesis because, while it specifies that there must be some intimate relation between remembering and epistemic success, it stops short of specifying that the epistemic success is *knowledge*. Though I will discuss other kinds of remembering in chapter 4, the focus of this chapter is fact remembering. The two generic formulations can be adapted for this purpose as follows:

(GET-P) A subject *S* remembers that *p* at *t* just in case she knows that *p* at *t* by operations of her faculty of memory.

(GET-P*) A subject *S* remembers that *p* at *t* just in case she has a justified true belief that *p* at *t* by operations of her faculty of memory.

Both theses enjoy *prima facie* plausibility, as many of our justified beliefs and many of the facts we know are things that we remember. For example, I am presently justified in believing that the Netherlands defeated Mexico 2-1 during the 2014 World Cup only because I remember that it did so. Of course, I know that the Netherlands defeated Mexico on the very same basis. Given this, there is some general methodological reason to prefer adopting the stronger of the two theses. Namely, it provides more theoretical mileage. The weak thesis characterizes the relationship between memory and justified beliefs. This is perfectly fine; however, it leaves a further important question unanswered – namely, what is the relationship between memory and knowledge. The strong thesis answers this question. In section 3.2, I develop this observation. In chapter 5, I offer a sustained defense of it.

3.2 Arguments in Favor of Strong G.E.T.

The argument in this section takes the form of a series of inferences to the best explanation. In particular, I present five phenomena that are better explained by the strong version of the generative epistemic theory than they are by the weak version of the theory.

Knowledge in memory contexts

First, the strong thesis provides the best explanation of our knowledge in memory contexts.[1] Specifically, it explains why "*S* remembers that *p*" is a fully satisfying answer to the question "How does *S* know that *p*?" For example, suppose that I report knowing that Pluto used to be classified as a planet. If you doubt me, you may go on to ask me how I know that fact. I fully respond to your inquiry simply by saying that I remember that Pluto used to be classified as a planet. Of course, I could (though need not) go on to speculate about where I learned this fact; however, given that I do actually remember that Pluto used to be classified as a planet, reporting as much explains how I know it.

Equally telling is the fact that challenging this explanation amounts to challenging whether or not one remembers. For example, suppose I report knowing that Berit visited the department in 2011. Again, when asked how I know this, I report that I remember that she visited in 2011. If you are aware of the fact that she did not visit in 2011 (you know that she visited in 2012), you are in a position to challenge my knowledge claim, but only by pointing out that I don't really remember what I think I remember.

[1] Ranalli (2014) and Cassam (2007) offer analogous discussions for perceptual contexts.

The strong epistemic thesis predicts these parallels. On that view, remembering something is a way of knowing it, and if one fails to know something, one surely does not remember it. The weak epistemic thesis does not directly predict or provide any explanation of these parallels because it simply says nothing about the relationship between remembering and knowing.

Felt inconsistencies[2]

Consider the following sentences:

1. 'Lara remembers that we should take a left on 1st but she doesn't know that we should take a left on 1st.'
2. 'Tim remembers that there was a big bald man on the corner, but he doesn't know that there was a big bald man on the corner.'
3. 'Jill remembered that Kristina's birthday is today, but she does not know that Kristina's birthday is today.'

All of these sentences *seem* inconsistent. An excellent explanation for why they might seem inconsistent is that they actually are inconsistent. And the strong version of generative epistemic theory explains why they would be inconsistent. If remembering that *p* is a species of knowing that *p*, then one cannot remember that *p* and also not know that *p*.[3]

An advocate of the weak theory might be tempted to argue that these sentences aren't really inconsistent, they merely give the appearance of being inconsistent because, ordinarily, when one remembers that *p*, one knows that *p*.

[2] This argument comes from Moon (2013: 2725-6).
[3] Bernecker (2010: 83) addresses similar tensions in first-person reports. It is not clear that his response is applicable to third person reports – see Moon (2013: 2724-5) for discussion.

This might provide a partial explanation of the data, but it is not clear that it is the best explanation. As Andrew Moon (2013: 2726) observed, the contingent correlation this response relies on would explain why the sentences might seem odd, but it would not clearly explain why they would seem *inconsistent*. I borrow his contrast case, "Suppose that Jordan and Kristina are, in ordinary cases, always together. One day, I tell a friend, 'Jordan was at the party, but Kristina was not.'" Given that they are ordinarily always together, one might find this assertion to be odd, and perhaps even suspect that it is false. However, there is nothing about the claim that would make it seem inconsistent. Consequently, a better explanation of this data is that sentences 1-3 really are inconsistent; this best is best explained by the truth of strong epistemic theory.

Contextual effects

A third observation is that memory and knowledge attributions appear to be sensitive to context in many of the same ways. Consider two cases adapted from Keith DeRose (1992: 913):

> *Bank Case A:* My partner and I are driving home on a Friday afternoon. We plan to stop at the bank on the way home to deposit our paychecks. But as we drive past the bank, we notice that the lines inside are very long, as they often are on Friday afternoons. Although we generally like to deposit our paychecks as soon as possible, it is not especially important in this case that they be deposited right away, so I suggest that we drive straight home and deposit our paychecks on Saturday morning. My partner says, "What if the bank is closed tomorrow? Lots of banks are closed on Saturdays." I reply, "No, I know this one is open on Saturdays. I was just there two weeks ago on Saturday. It's open until noon."
>
> *Bank Case B*: My partner and I drive past the bank on a Friday afternoon, as in Case A, and notice the long lines. I again suggest that we deposit our paychecks on Saturday morning, explaining that I was at the bank on Saturday morning only two weeks ago and discovered that it was open until noon. But in this case, we have just written a very large and very important check. If our paychecks are not deposited into our checking account before Monday morning, the important check we wrote will bounce, leaving us in a *very* bad situation. Again my partner says, "What if the bank is closed tomorrow? Lots of banks are closed on Saturdays." Though I am just

> as confident that the bank is open on Saturdays as I was in Case A, I confess that I don't *really know* that the bank is open on Saturdays.

Some take these cases to show that knowledge attributions are sensitive to context.[4] Specifically, an attribution of the form, "Steven knows that the bank is open on Saturdays" is plausibly true in the context described in Case A but false in the context described in Case B. The general thought is that when the stakes are higher, so are the standards for knowing.

Suppose that this assessment of the cases is plausible. Notice that replacing "knows" with "remembers" in the relevant places would merit the exact same judgments. Specifically, an attribution of the form, "Steven remembers that the bank is open on Saturdays" is plausibly true in the analogue of Case A and false in the analogue of Case B. The best explanation for this parallel pattern of judgments is that remembering is sensitive to context in the same way that knowledge is because remembering is simply a *species* of knowledge. And again, given that the weak version of the generative epistemic theory says nothing (directly) about the relationship between knowledge and memory, it does not directly explain this phenomenon.

Purposes of reports

Along similar lines, knowledge and memory reports seem to share at least one significant purpose in our linguistic practices. Specifically, '*S* knows that *p*' and '*S* remembers that *p*' are both used to indicate that *S* is a good source of information

[4] See Cohen (1999), DeRose (2001), and Stanley (2004) for more discussion of these kinds of cases.

with respect to *p*.[5] So, for example, one indicates that Jen is the right person to talk to about the creators of a certain album by reporting that either, "Jen knows who recorded, 'Transmissions from the Satellite Heart'" or, "Jen remembers who recorded 'Transmissions from the Satellite Heart'".

Again, the most natural explanation of this parallel is that remembering reports indicate that subjects are good sources of information because knowledge reports do, and remembering is just a species of knowledge.

Norms of assertion

Finally, some take knowledge to be the norm of assertion. That is, they accept the claim that one is only permitted to assert that *p,* if one knows that *p*.[6] For those who accept this thesis, the strong theory of generative epistemic remembering provides the best explanation for why it is permissible to assert that *p* if you remember that *p*. To continue the example, if Jen remembers that the Flaming Lips recorded "Transmissions from the Satellite Heart", she can assert that they did so when you ask her about it.[7]

The upshot of this discussion is that the best available theory of remembering is not just a generative epistemic one, but it is a particularly strong

[5] See Craig (1990), Reynolds (2002), Henderson (2011, 2009) and Henderson and Horgan (forthcoming) on this purpose of knowledge attributions.

[6] See e.g. Garcia-Carpintero (2004), Hawthorne (2004), DeRose (2002), and Williamson (2000: 243). See Slote (1979: 185), Unger (1975: 253-70) and Moore (1912) for some predecessors.

[7]Given that many associate the claim that knowledge is the norm of assertion with Williamson (2000), it is worth noting that one could appeal to material in his framework to make at least two further arguments for the strong version of the epistemic theory of remembering. First, if Williamson is right that knowledge is the most general factive stative attitude, the strong epistemic theory of remembering would follow (given the quite plausible assumption that remembering is a factive, stative (non-complex) attitude). Second, one could combine Williamson's view (2000: 184-207) that one can only (appropriately) use *p* as a premise in inferential reasoning if one knows that p, with the observation that we typically have to remember the premises we use in inferential reasoning to conclude that remembering that *p* is sufficient for knowing that *p*.

version of generative epistemic theory. Factual remembering is not merely an epistemic success; it is a paradigm of epistemic success – it is a species of factual knowledge. Namely, it is the knowledge one gets via operations of one's faculty of memory. Having made the case for this thesis, I turn next to describing some initial payoffs of the theory.

3.3 G.E.T. and the Troublesome Cases

In chapter 2, we saw ten cases that challenged causal theories of remembering in a variety of ways. In this section, I briefly compare the causal theory's assessment of the cases with the generative epistemic theory's assessment – showing that the latter is more in line with our intuitive judgments about the cases, and provides a better explanation of why they are, or are not, instances of remembering that *p*.

Case 1 – Conceptual Enrichment: Bernecker's causal theory denies that certain cases involving conceptual enrichment can be instances of remembering that *p* because those cases fail to satisfy an entailment condition. This is unfortunate because it certainly seems possible to, e.g. remember that one went to a Catholic wedding by precisifying one's prior belief that one went to a wedding by appropriately adding the concept CATHOLIC. This is so even though the former belief does not entail the latter belief. The generative epistemic theory developed here fairs much better in such cases because it can be considerably more flexible about how one can add to one's knowledge. Entailment is one but not the only logical relation that can extend knowledge.

Case 2 – Tense: Similar points apply to cases of remembering that involve tense. That one's car is parked in Lot C at t_1 does not entail that one's car is parked

in Lot C at t_2. Consequently, the causal theory (as formulated) must deny that one can remember where one's car is parked upon returning home from vacation. The generative epistemic theory has no problem with such cases for it can allow that one knows things that are not *entailed* by anything else they believe.

Case 3 – Condensed Memories: The causal theory is also challenged by certain versions of condensed memory because it is not clear that any unique series of memory traces need strictly be an inus condition for the remembering state. Though the generative epistemic theory of remembering presupposes that remembering states are the end product of some or other causal processes, it allows for a plurality of candidate causal chains (restricted only to those that can sustain knowledge). Consequently, it faces no problem from condensed memory cases.

Case 4 – Hypnotic suggestions: The causal theory and the generative epistemic theory both rightly conclude that one cannot remember that *p* on the basis of having a hypnotist unknowingly plant a true belief that *p* in one's mind. The causal theorist concludes this on the basis of the intuition that there exists no non-deviant causal connections between the circumstances that make *p* true and one's relevant belief. The epistemic theorist does better by explaining *why* the lack of certain causal connections is a bad thing – the lucky nature of the resulting true belief impugns its status as knowledge.

Case 5 – Testimony: The causal theorist must weaken her theory to make sense of cases in which remembering that *p* involves a testimonial chain (testimonial chains do not themselves go only by way of memory traces). The generative epistemic theory has no problem accounting for such instances of

remembering because any plausible account of memory processing recognizes that testimonial information is among the information that memory processes generally incorporate into subsequent representations.

Case 6 – Unjustified memory beliefs: The causal theory fails to explain why unjustified true memory beliefs cannot be instances of remembering, without making an appeal to some epistemic resources.[8] The generative epistemic theory has a fully satisfying explanation. Since remembering is a species of knowledge, and knowledge (plausibly) requires justification, unjustified memory beliefs cannot be instances of remembering that *p*.

Case 7 – Lucky memory beliefs: Similarly, the causal theory fails to explain why a lucky memory belief that *p* might fail to be an instance of remembering that *p*.[9] The generative epistemic theory has a nice explanation of why certain lucky memory beliefs aren't instances of remembering (while also explaining why certain others are instances of remembering). If the luck at issue is epistemic luck, one cannot remember that *p*, for epistemic luck undermines knowledge; if the luck is consistent with knowledge, one can remember that *p*.[10]

Case 8 – Alternative life forms: Given the ordinary conception of memory traces, the causal theory is inconsistent with the possibility that there might be creatures that are biologically very different than humans, and yet do remember things. To accommodate this possibility, the causal theorist must sacrifice some of

[8] Bernecker (2010: 75-83) goes so far as to conclude that unjustified memory beliefs are instances of remembering so long as one's memory processes are operating normally.
[9] Again Bernecker (2010: 73-74) goes so far as to allow that lucky memory beliefs are instances of remembering that *p*.
[10] More anon in chapter 5.

the theoretical appeal of the causal theory by offering a revised, and a more abstract, conception of memory traces. The generative epistemic theory has no specific commitments regarding the physical processes that enable remembering. Consequently, it provides an account of remembering for a variety of possible creatures.

Cases 9 and 10 – Silicon chip replacement and downloaded memories: The same points apply to the possibility that neuroscientists could sustain one's mnemonic capacities by replacing dying neurons with silicon chips or by briefly downloading information from one's neural networks onto digital networks. The generative epistemic theory of remembering is consistent with a variety of different kinds of memory operations, whether biological, prosthetic, or otherwise. It requires only that the processes enable one to know that *p* at *t*.[11]

The upshot of this discussion is that a strong generative epistemic theory of remembering (hereafter the 'generative epistemic theory' unless otherwise noted) (i) is the most motivated epistemic theory of remembering, and (ii) outperforms its most prominent non-epistemic competitor (the causal theory) with respect to a variety of phenomena. Consequently, we have good reason to conclude that memory is a source, and remembering is a species, of knowledge.

3.4 Knowing by Operations of Memory

The generative epistemic theory of fact remembering has two crucial components. First, it makes knowing that *p* a necessary condition for remembering that *p*. Second, it specifies that the relevant processes that support one's knowing of *p* must

[11] This level of flexibility presupposes that some versions of extended mind theses are plausible, but I do not take this to be a substantive cost of the view.

be memory processes. Since these are taken to collectively suffice for remembering that *p*, one might be tempted to ask what it is to "know something by operations of one's memory faculty."

Strictly speaking, the generative epistemic theory is silent on this issue, and for good reason. First, even if we restrict the investigation to the operation of human memory faculties, there is little reason to think that there is one unique processing system or set of principles that govern all instances of remembering (cf. the discussion of memory systems in chapter 1). Moreover, even if there were, those processing mechanisms would only contingently support remembering. A theory of remembering should not rule out the possibility of different mechanisms either evolving or being designed to enable the very same psychological phenomena. Second, it would also require a characterization of propositional knowledge and such a difficult project is well beyond the scope of a theory of remembering.

One might object that the plausibility of the generative epistemic theory depends on details about the theory of knowledge. After all, given what we do understand about *typical* human memory processing, much of what we seem to remember would fail to qualify as remembering in the context of a particularly demanding theory of knowledge (e.g. one that requires Cartesian level certainty).

Though certainly true, this is no objection to the view. Rather, it is quite plausible that what counts as an instance of remembering may depend, in part, on one's broader theoretical commitments. That said, the generative epistemic theory of remembering is perfectly in line with what appear to be very natural desiderata for a theory of knowledge. Namely, the theory of knowledge should not be so

permissive as to count every representation that emerges from one's cognitive systems (including memory) as knowledge. Nor should it be so restrictive as to lead to general skepticism (including skepticism about memory knowledge).

This level of generality and flexibility might lead one to a second objection to the generative epistemic theory. Specifically, one might claim that the view fails to be an informative theory of remembering because it is circular. It analyzes remembering in terms of knowing by operations of memory and it offers no reductive characterization of memory operations.

There are several things to say in response to such a worry. First, it is worth observing that over the course of the last 50 or so years, philosophers have come to recognize that not all informative accounts of phenomena must take the form of reductive analyses. Moreover, the kind of circularity at issue in this analysis is not particularly vicious. The view does not define remembering in terms of memory with the expectation that memory is to be defined in terms of remembering (as such). While it carries no commitment to any particular memory mechanisms, it is worth observing that typical human memory processes are examined and understood in a variety of ways. The theory can simply defer to (or borrow from) the methodology of natural science as appropriate in each case.

For example, psychologists examine the properties of typical human memory processes by looking at a combination of behavioral and biological data. Roughly, they identify paradigmatic instances of a target behavior; they then explore the neurological basis for such behaviors by finding correlations between said behavior and neurological activity. They can then examine the roles said neurological activity

may have in other target behaviors, and so on and so forth. Of course, this relies on a background assumption that one can identify paradigmatic instances of remembering, but one need no theory of remembering to accomplish that task.

As they uncover additional associations and dissociations between different neurological activities and behaviors it becomes possible to talk about memory processes (those associated with paradigmatic instances of remembering) in terms of behaviors that are not themselves instances of remembering. For example, many memory processes are now known to also play crucial roles in other behaviors such as imagining, reasoning, counterfactual reasoning, planning, predicting, and so forth. Consequently, the analysis of remembering may involve a circle of sorts, but it is not so small as to render the account uninformative.

Though the memory processes uncovered by empirical psychology are not strictly necessary conditions for remembering, it may still be useful to review the basics of the mechanisms that support remembering in some cases in which a normal (hypothetical) human subjects know that *p* by operations of their memory faculties.

Recall the basic principles of memory introduced in chapter 2 (pp. 50-53):

(i) Selection: Only certain incoming stimuli are selected for encoding.

(ii) Abstraction: The meaning of a message is abstracted from the syntactic and lexical features of the message.

(iii) Interpretation: Relevant prior knowledge is invoked.

(iv) Integration: A holistic representation is formed from the products of selection, abstraction, and interpretation processes.

(v) Reconstruction: During retrieval, whatever information was selected for representation and is still accessible is used, together with general knowledge, (roughly) to generate a representation of what must have happened.[12]

We can illustrate how these principles govern typical human subjects coming to *know* by operations of memory with two familiar examples. First, consider again the case used to illustrate how a subject remembers a car accident and certain facts about it. Suppose that a witness undergoes a visual experience of a car accident at t_1. Although the experience is tremendously rich, only some of the visual stimuli are selected (e.g. the shapes and colors of the two involved vehicles, and so forth) by her cognitive systems for further processing. The selected information is abstracted and interpreted using available prior knowledge and all of this is integrated into a holistic representation of the event (e.g. she comes to have a representation of an accident event involving one blue and one red car). Suppose that the subject is later (at t_2) shown images of the accident wreckage, selects, abstracts, and interprets the information available in that experience to form an additional representation of the seen images.

[12] Two points are worth making here. First, the "must" in this principle should be read as something like an inference to the best explanation given all the available information. Some prefer to formulate it with "might". However, I think this would be somewhat misleading. It does not follow from the fact that memory processes construct a representation that is largely predictive that, in remembering, one only predicts what might have been the case. See DeBrigard (2014) for a differing point of view and chapter 5 for some further discussion. Second, this set of principles is not the only set offered in the psychological literature. For example, Surprenant and Neath (2009) offer a set of principles that also emphasizes the role of cues as facilitators of retrieval, and effectively does a bit more to predict the conditions under which one might fail to remember something. As far as I can tell, these proposed sets of memory principles are largely compatible and nothing of philosophical interest will turn on any differences they might have in detail.

Now suppose that she is asked the following (at t_3): "What directions were the respective vehicles heading at the time of the accident? Which vehicle initiated the impact? And what was the point of impact?" This prompts factual remembering. Her memory processes piece the information that had previously been encoded and interpreted at times t_1 and t_2 (or as much of it as remains available) together with her other general knowledge about physical interactions (e.g. she knows that an object that has a dent on the left front fender has been impacted from the (relative) left front direction, etc.) to *construct* representation of what happened. As a result, she comes to *know* that the blue car was headed north; the red car was headed south; and the red car crossed the center lane impacting the blue car at the left front fender by virtue of operations of her faculty of memory. That is, according to the generative epistemic theory of remembering, she comes to remember that the blue car was headed north, and so on and so forth.

Similarly for a subject's operations of memory enabling her to come to know (for the first time) that she had been to a Catholic wedding. In that case, whatever information is left from her prior representation of the wedding event, is integrated with information she has subsequently acquired about Catholicism, to form a novel bit of knowledge.

The upshot of this discussion is that the generative epistemic theory commits to no particular account of knowing by operations of one's faculty of memory. There probably is no univocal account of this even for normal human subjects, but there certainly is no single account that is general enough to be considered part of a

general theory of fact remembering. In short, the generative epistemic theory of remembering is a non-reductive, yet still informative account of fact remembering.

3.5 Conclusion

In this chapter, I characterized and argued for a strong version of generative epistemic theory of fact remembering; namely, a subject *S* remembers that *p* at *t* iff *S* knows that *p* at *t* by operations of her faculty of memory. We saw that this theory outperformed its central competitor theory – the non-epistemic causal theory of fact remembering – on a variety of cases. We also saw that, though a non-reductive account of fact remembering, the theory remains informative and compelling.

Chapter 4 G.E.T Objects and Events

There are two central goals for this chapter. First, I will argue that the generative epistemic framework introduced in chapter three can be extended to at least some non-propositional remembering states; namely, the remembering of objects and some events. As the aim is to simply show that the theory can be extended, I will gloss over some interesting differences between the remembering of certain kinds of events (usually first personal) and the remembering of objects. Consequently, the discussion should not be viewed as a complete account of non-propositional remembering states. The second goal is to examine two related phenomena – visual hallucinations and visual imaginings of real things (e.g. persons or prior events). I argue that both are a species of unusual remembering. The chapter is organized as follows.

In section 4.1, I discuss four preliminary objections to the very idea that non-propositional remembering states are amenable to an epistemic treatment. I argue that none of these objections are deep, i.e. they do not challenge the spirit of the generative epistemic theory. In section 4.2, I formulate a generative epistemic theory for objects and at least some events. I then describe one way one might develop the proposal. In section 4.3, I shift the discussion to visual hallucinations (and imaginings) of real things. I argue that current accounts of such experiences are inadequate. In section 4.4, I argue that these states are an unusual species of perceptual remembering. And in section 4.5, I defend this thesis against objections.

4.1 Preliminary Objections

Some have thought that an epistemic theory of non-propositional remembering is simply a nonstarter. The general thought is that subjects cannot stand in the requisite kind of epistemic relations to objects and events, i.e. they are not the right kinds of things to be known. I will consider four arguments.

Error

Alex Byrne (2010:20) has argued that event remembering cannot be a species of knowledge because it permits error in ways that knowledge does not.[1] He offers the following case to illustrate:

> Suppose one seems to recall being sober at the party, wearing a pink tie, talking to McX, and so on. In fact, one was drunk, tieless, and McX never turned up. It is consistent with misrembering all these details that one is genuinely recalling the party. One may misremember practically everything (including misremembering the party as a faculty meeting), and yet still be recalling the party. (Byrne 2010: 20)

The central thought is that propositional knowledge admits of no error – you only know that *p*, if *p*. This case shows that remembering events is a more complicated matter. You can get many things about the event wrong and yet still remember it.

Byrne's reasoning supports the conclusion that event remembering is not a species of *propositional* knowledge. It does not follow that remembering an event is not a species of any kind of knowledge. In particular, it remains perfectly plausible that remembering objects or events is a species of *object-involving* knowledge.[2] After all, our ordinary object-involving knowledge does seem to typically permit error in the same ways that our object and event remembering does. For example, one can both remember and know Alex, despite being wrong about whether, e.g. he has blue eyes, stands 6' tall, or so forth. So, the fact that event and object remembering permits error does not directly challenge an epistemic theory of object and event remembering.

Mike Martin (2001) offers three further arguments against an epistemic theory of event or object remembering: (i) an argument from grammar; (ii) an argument from our

[1] Though he formulated the argument for events, the same points could be applied to objects as well.
[2] Here and throughout this chapter, I will use "objectual" and "object-involving" interchangeably. Strictly speaking, objectual knowledge is probably a subset of object-involving knowledge as, arguably, some propositional knowledge is also object-involving, but this will not matter here since the focus of this chapter is on objectual knowledge and objectual/event remembering.

linguistic preferences; and (iii) argument from the temporal boundedness of objects and events.

Grammar

First, he argues that we cannot even conceive of event remembering as preserved knowledge for one cannot have known what one now recalls (2001: 264). To be clear, Martin is not merely attacking the preservative nature of remembering; if successful, his argument challenges both preservative and generative epistemic theories.

He offers the following as evidence for his conclusion:

1) Mary remembers John falling asleep in the talk.
2) *Mary knows/knew John falling asleep in the talk.
3) Jo recalls being inoculated for smallpox.
4) *Jo knows/knew being inoculated for smallpox.

The suggestion is that, since sentences 1 and 3 are perfectly ordinary reports of event remembering, and sentences 2 and 4 are, "simply not English", we should conclude that event remembering is not a species of knowledge.[3]

The problem, as Byrne (2010: 20) points out, is that the fact that expression A cannot be substituted for expression B, preserving grammaticality, does not guarantee that A and B differ in semantic value (at least not when A and B are of different syntactic

[3] Martin ultimately concludes that event/object remembering is "retained acquaintance" or "retained apprehension". It is not clear to me that this is inconsistent with the spirit of an epistemic theory as it would be natural to think that one has retained *knowledge* by acquaintance in these cases. Martin himself frames the issue as follows. To have known *o* one must have apprehended *o*, but to continue knowing *o*, one must have the possibility of further apprehension of *o*, and we are not in any position to re-apprehend token events, and many objects. He does, however, view "apprehension" as a kind of distinctive cognitive success. It appears to be the episodic counterpart to knowledge and both of these cognitive successes are species of the more basic "cognitive contact." Unfortunately, Martin does not provide a thorough characterization of "cognitive contact". More on the conditions Martin takes to be necessary for knowing *o* at a time below.

categories).[4] In short, it simply isn't clear what we are supposed to conclude from Martin's argument. One plausible suggestions is that, if one has objectual knowledge of an event by remembering it, we will not report it with direct object constructions of the form, 'S knows *o*' where '*o*' is a noun or gerund phrase. This, of course, is compatible with an epistemic theory of event remembering.

Temporal boundedness

Martin builds his second argument around two cases (2001: 263).

Case 1

> Spurred to think about my schooldays, I may recall several of my classmates. If I have lost touch with them, and not concerned myself in the interval with their progress in the world, then I might comment that I did know S or X but that I know them no longer. On the other hand, I can certainly correctly report myself as remembering them.

Case 2

> I cannot now know my maternal grandmother, for she died twenty years ago, but I did know her, in contrast to my paternal grandmother who died when I was first born. On the other hand I can certainly remember my maternal grandmother, and cannot remember my paternal one.

Martin takes these cases to show that one can remember *o* at *t* without knowing *o* at *t*. The explanation for this is that objects exist within time and can come to be, alter, and cease to exist. Consequently, knowing individuals requires having encountered them, but continuing to know them also requires tracking their subsequent changes. When you don't track such changes, though you can be said to have known them, you can no longer be said to know them.

There are several responses to make to such an argument. First, it is not clear that continuing to know something requires tracking its subsequent changes. For example, though Martin takes the first case to have shown that he does not know now anyone he

[4] See also King (2007, ch. 6).

went to school with (as a child), this is not obviously true; or at least it need not be true for anyone in an analogous situation. There may be a sense in which one does not know one's childhood classmates *well*. But it is quite possible for one to still know them in another sense. To see this, suppose that you encounter a childhood schoolmate as an adult *and remember* them as such upon sight. You could truthfully proclaim, "I know you!" And the truth of your knowledge claim would come, not from your present perception of the individual, but from your remembering them (one cannot assert "I know you!" upon seeing someone whom one does not remember.) Moreover, it seems to true to say that one knows the individual in this sense, even if one did not actually encounter and remember them, but if one would remember them, if they were to encounter them. So it seems Martin has drawn the wrong conclusion in this first case.

In contrast, he seems to have drawn the right conclusion in the second case. He cannot now know his maternal grandmother. However, it also seems that there is a way to understand this case that is consistent with (the spirit of) an epistemic account of object remembering. Here is a first pass at an attempted explanation: Martin cannot know his grandmother (qua existing person) by remembering her because he has no existing grandmother. Of course, he cannot remember his grandmother (qua existing person) either, for there is no such thing. He can, however, know his grandmother (qua previously existing person) by remembering her for she did previously exist.

To summarize, there are several potential things at issue in each case. In the first case, Martin can be said to know both, the existing individual by remembering him, and the past individual by remembering him. In the second case, Martin can be said to remember the past individual but not the present individual. The difference is simply that there is no

present individual to know or remember in the second case. This suggests that actually tracking changes in a thing is not a necessary condition for continued knowing of the thing. Rather, continued existence of the thing is the necessary condition for continued knowledge of the thing.

There is, of course, a somewhat parallel phenomenon in propositional knowledge and fact remembering. I argued in chapters 2 and 3, that having parked my car in Lot C at t_1, I can remember and know where my car *is* parked at t_2, upon returning from holiday. Of course, I can only remember and know this if my car is still parked in Lot C at t_2. In neither the fact-remembering/knowing nor the object-remembering/knowing cases is actually tracking changes (in features or states of affairs) requisite for knowing those respective objects or facts. What does seem to be crucial is whether, at t_2, a certain state of affairs obtains or a certain object exists. Given this parallel, Martin's second argument does not seem to directly challenge (the spirit of) an epistemic theory of object remembering.

Linguistic reticence

The last argument Martin gives is unique to event remembering and has the form of an appeal to our linguistic practices. The problem, as he sees it, is that we are extremely reluctant to talk about knowing events. So, for example, while one can know Paris, one cannot be said to know the Vietnam War, one can only witness it.

This argument can be treated in much the same way as Martin's first argument. It is simply not clear what it is supposed to have shown. Indeed, one might think that witnessing an event is a way of coming to get a particularly intimate knowledge of it. At most, the conclusion these kinds of arguments seems to license is that, if one has object-involving knowledge of an event by remembering it, we will not report it with direct object

constructions of the form, 'S knows *o*' where '*o*' is a noun or gerund phrase. And again, this is compatible with an epistemic theory of event remembering.

4.2 G.E.T. Object and Events

None of the arguments in section 4.2 rule out the possibility of providing a generative epistemic account of object and event remembering. They do, however, highlight that we must be careful about how to best formulate such a proposal. The following is a natural proposal:

(GET-O) A subject *S* remembers an event or object *o* at t_1 just in case she has object-involving knowledge of *o* at t_1 by operations of her faculty of memory.

Before proceeding, I must address a tempting, but ultimately misguided, objection to the proposal. Martin (2001: 265 n. 11) briefly considers and dismisses the possibility of formulating the knowledge one has in virtue of event/object remembering using a 'knowledge of' locution. His objection is that the 'knowledge of' locution picks out the kind of knowledge that, e.g. a historian has about the Vietnam War in virtue of knowing *the subject matter* – it refers to knowing a set of facts about it.

He is certainly right that this is the most familiar use of the 'knowledge of' locution. However, dismissing the construction on this basis would be premature. The generative epistemic theory does not simply require that one have knowledge of the event in any sense of the locution. It requires that one have *object-involving* knowledge of the event, *by operations of the faculty of memory*. This combination of requirements is not obviously met by cases in which one simply knows much about a subject matter. Knowing about a subject matter does not obviously involve object-involving knowledge at all; nor need it come by operations of the faculty of memory.

Given these qualifications, there are two plausible sources of 'knowledge of *o*' in the relevant sense. The most familiar source is direct perception of *o,* as in when one sees an object or event. A secondary, but no less significant, source is mediated perception of *o*, as in when one encounters a mediated representation of *o*, e.g. recordings of Obama's inauguration, the Challenger disaster, Reverend Martin Luther King Jr., the assassination attempt on Reagan, the successful assassination of JFK, and so forth. (*cf*. the distinction between autobiographical and non-autobiographical event remembering in chapter 1 (pp. 9-10)).

So, we have a *prima facie* plausible account of object and event remembering that parallels the account of fact remembering developed in chapter three. As should be expected, the account has two crucial components. It makes object-involving knowledge of *o* a necessary condition for remembering *o*. It also specifies that the relevant processes that support one's knowledge must be memory processes. And as expected, one might be tempted to ask the analogous question here as well. Namely, what it is to have object-involving knowledge of *o* by operations of one's memory faculty?

Just as before, the generative epistemic theory is silent on this issue, and for good reason. Again, there is little reason to think that there is one unique processing system or set of principles that govern all instances of remembering (cf. the discussion of memory systems in chapter 1). Moreover, even if there were, those processing mechanisms would only contingently support remembering and a theory of remembering should not rule out the possibility of different mechanisms either evolving or being designed to enable the very same psychological phenomena. And, in this case, it would also require a characterization of object-involving knowledge – a project well beyond the scope of this dissertation.

That said, there is a familiar development of object-involving knowledge that is worth at least briefly mentioning. Namely, one might develop the notion of object-involving knowledge in terms of knowledge by acquaintance (loosely modeled on Russell's 1912 notion).[5]

There is a natural model for doing so that is available in the literature on perception (cf. Tye (2009: 96-102)). In perception, one might say that one has knowledge by acquaintance of *o* by seeing *o*, without seeing that *o* is any particular way (and so without knowing any particular truth about it). Analogously, we might say that one can have object-involving knowledge of *o,* qua having knowledge by acquaintance of *o,* by remembering it without remembering any particular facts about it.

In both cases, the mental state one occupies in seeing/remembering is undeniably an epistemically enabling in that it puts one in a position to know facts about the thing (Tye, 2009: 98). Of course, exactly how we are to understand what it is to be in a position to know facts about a thing is itself a difficult matter that cannot be properly pursued here.[6] Moreover, given the distinction between autobiographical and non-autobiographical object and event remembering, it is clear that 'acquaintance' will be a considerably more liberal relation than either Russell or Tye respectively endorse.

Nevertheless, we have seen that the generative epistemic theory of remembering can be extended to address at least some cases of object and event remembering. Moreover, there is at least one promising strategy for developing the proposal beyond the commitments that have been adopted here.

---

[5] Doing so makes the disagreement between the proposal I have made here and Martin's preferred characterization quite minimal.

[6] See Tye, 2009: 96-102, for some discussion.

4.3 Hallucinating Real Things

In Act II, Scene I of Shakespeare's *Macbeth*, the titular character has a visual experience as of a dagger floating before him. To Macbeth it was as though his experience was of a particular worldly object. It was not, and much has been done to make sense of this kind of phenomenon.[7] Here, I focus on another kind of hallucination. As Macbeth descends deeper into madness, he has a visual experience as of his former comrade Banquo, whom he had recently betrayed. Unlike the dagger-hallucination, this visual experience *was* of a particular external world object—namely, Banquo.[8] This kind of object-involving hallucinatory experience has unfortunately only received passing attention in the philosophical literature.[9] These hallucinations connect the subject to external world objects and so should not be understood in precisely the same way as their object-independent counterparts.[10] At the same time, they provide a distinctively *sensory* kind of connection to those particulars, so they should not be understood in the very same way as object-involving thoughts are in general. Something else is needed.

---

[7]See e.g. Crane (2011a), Johnston (2004), Smith (2002) Smith (1983), and especially Macpherson (2013) for discussion of the role this kind of hallucination has had in shaping much of contemporary philosophy of perception over the last hundred or so years.

[8] Here I invite the reader to suppose that the fiction is fact but this is inessential as I use fictional cases only for simplicity of illustration. Corresponding case studies of both object-involving and object-independent hallucinations are readily available in the psychological literature. See e.g. Sachs (2012: chapters 1 and 13), Vignal *et al*. (2007), and Carroll and Carroll (2005) for discussion. While the empirical literature does not generally focus on the importance of the distinction between object-involving and object-independent hallucinations, the literature is largely compatible with (and, in some cases, directly supports) the proposal I develop in section two.

[9] Johnston (2004: 129-132) provides the most sustained discussion of the phenomenon; Pryor (2007) and McLaughlin (1987: 116-117) each dedicate roughly two paragraphs to object-involving hallucination; Dorsch (2010 n. 1) and Tye (2009 n. 2) mention it in footnotes to their discussion of hallucinatory experiences; and although Smith (2002:266) appears to deny that such experiences are properly understood to be hallucinations at all (see also Tye (Ibid.) for a similar position), he does recognize the phenomena as a genuine psychological possibility and his discussion of it generally follows McLaughlin's earlier approach.

[10]Dorsch (2010) maintains that the relation is an *epistemic* one but says nothing else about the matter. I agree with him on this point, but I think more can, and should, be said about exactly what kind of epistemic relation object-involving hallucination provides.

I propose that these experiences are an unusual species of perceptual remembering. Here I define *perceptual remembering* as the remembering of objects of perception in a distinctively visual manner—a manner that involves the (re)presentation of external objects or events rather than propositions about them.[11] I begin here by criticizing the most developed account of these phenomena that is available in the literature. In section 4.4, I develop the claim that such experiences are best understood as instances of perceptual remembering. And in section 4.5, I defend the proposal against objections to the effect that object-involving hallucinations cannot be instances of remembering because they fail to share some essential feature or features with memory.

Hallucinations like the one Macbeth had of Banquo represent real-world objects, events, and qualities. However, as hallucinations, they cannot themselves be the original source of such representations.[12] Given this, it has generally been recognized that they must be "anchored" to the particular objects they represent via some other mental state(s) of the hallucinating subject. And so a crucial task for one aiming to understand object-involving hallucination is to provide a satisfactory characterization of the anchoring

[11] It is important that one not confuse perceptual remembering, as I understand it here, with the more specific category of 'episodic memory' as it is defined in, e.g. Tulving (1972) or Martin (2001). Not all perceptual remembering is episodic. That is, not all perceptual remembering: (i) answers the retrieval query, "What did you do at time T in place P?" (*cf*. Hasselmo (2012:2)), (ii) is naturally reported using constructions of the form '*S* remembers [x] *f*–ing' (*cf.* Martin (2001: 261)), and (iii) involves "mental time travel" as characterized by Tulving (1983). See Tulving's somewhat facetious (2007), and Schacter and Tulving's (1994) for discussions of the empirically recognized categories of memory and memory systems. It is also worth noting that my use of 'perceptual remembering' differs from Malcolm's (1963) use of 'perceptual memory'. The latter corresponds more closely to 'episodic memory' as it is generally described.

[12]Though the proper characterization of hallucinatory experience is debatable (see MacPherson (2013) for discussion), I will understand perceptual hallucination to be an experiential state in which, contrary to internal appearances, there are no external world objects being *perceived*. This can be contrasted with other non-veridical perceptual states such as illusion in which there is an object perceived, but the perceiver experiences it as having properties it lacks (*cf*. Huemer (2007, 2001: 127), Johnston (2006: 269), and Brain (1962: 828)).

relation or relations. In the rest of this section, I argue that lack of serious attention to this kind of hallucination has left us without any such satisfactory characterization.

The general consensus has been that, when hallucinating, a subject can hallucinate an external world object if she can deploy some antecedently obtained singular or object-directed representational element (*cf*. McLaughlin (1987: 116), Pryor (2004), and Johnston (2004: 129-132)). And it is plausible that having a singular or object-directed representational element about a particular external world object will depend in some crucial way on memory (*cf*. McLaughlin (Ibid.) and Smith (2002: 266)).[13] However, much more must be said, as the same is true of other mental states as well, e.g. thoughts about particular objects. Mark Johnston (2004:132-3) has offered the most developed attempt at an account of the phenomenon as follows:[14]

A subject *S* visually hallucinates an external world particular *o* when

1. *S* is aware of some visual profile (a complex of visual properties) *p*.
2. *S* is not aware of instantiations of *p* in the scene before her eyes.
3. One or more of the following conditions obtains:
    a. *p* strikes *S* as being *o*;
    b. *S*'s visual awareness of *p* is caused in the right way by either,
        i. an earlier perception of *o*, or

[13] Two quick clarifications are appropriate here. First, this assumes a causal/acquaintance based understanding of singular thought/object-directedness as featured in, e.g. Bach (2010, 1987), Böer and Lycan (1986), Brewer (1999), Donnellan (1979), Evans (1982), Recanati (2010, 1993), Lewis (1979), Kaplan (1989a, b), Salmon (1986), Soames (2005, 2003), Pryor (2007), and Tye (forthcoming, 2009)). Second, I am not assuming that "singular" and "object-directed" are equivalent notions. I will, however, use them interchangeably to remain neutral on the relationship between singularity and object-directedness. See Hawthorne and Manley (2012), Crane (2011b), Armstrong and Stanley (2010), Jeshion (2010); Sainsbury (2005); and McKinsey (2009) for discussions of the relationship between object-directedness and singular thought/singular propositions.

[14] To be fair to Johnston, he has offered this proposal up as a "first pass" at an account of object-involving hallucination. That said, even in such a cursory form, his discussion does seem to capture the general line of thinking about the matter in the philosophical literature to date.

ii. a thought to the effect that *o* is such and such.

While Johnston's account is couched in terms that may not be shared by others working on the topic, his proposal is a natural development of the general line others have taken.[15] It is, however, inadequate. Clauses 1 and 2 specify that the subject is undergoing a visual hallucination and can be set aside for the purposes of this discussion. Clause 3 is meant to characterize the anchoring relation(s) and it will be the focus of the rest of this section.

The clause itself is divided into three conditions, the obtaining of any of which is meant to suffice to make an hallucination object-involving. I will argue that, as formulated, none of the three guarantees that the subject has an object-involving hallucinatory experience. Consequently, Johnston's proposal fails to adequately characterize the phenomenon of interest and the general approach must be revised.

According to the first anchoring clause (3a), a subject can hallucinate an external world particular simply by having an uninstantiated visual profile "strike" her as being that particular. So, Banquo was the object of Macbeth's hallucination simply because, to Macbeth, it seemed to be an experience of Banquo. The problem is that although it is often met when one is hallucinating a particular object, this condition is neither necessary nor sufficient for a hallucination's being object-involving.[16] For our purposes, its status as a

[15] The notion of a "visual profile" is perhaps best paired with the "direct realist" account of perception that Johnston (2006, 2004) and Sosa (2011, 2007, 1996) have developed. As a result, it may be rejected by those favoring other approaches to perception. Otherwise, the proposal is a natural, and the most worked out, development of McLaughlin's (1987:116) earlier idea that one can hallucinate an object if one can form a thought about it (see also Pryor 2007). And it is a reasonable development of Smith's (2002: 266) thought that such experiences depend crucially on memory.

[16] It should be no surprise that the visual profile in an hallucination of an object *o* might strike the hallucinating subject as being *o.* Hallucinations, as philosophers understand them, can be indistinguishable from veridical perceptions from the inside. So, if one is hallucinating *o*, it is likely that the visual experience

sufficient condition is all that matters and it is *not* sufficient because "striking", in Johnston's sense, amounts to nothing more than the subject purporting to "identify" the apparent object experienced as being *o*. And to "identify" an apparent object as *o*, a subject need only be able to deploy an individual concept of *o* in connection with her present visual experience. If one has the ability to deploy an individual concept of *o* at all, it is plausible that such a concept could, in principle, get linked to *any* visual profile such that that profile strikes the subject as being *o*.

This could not, by itself, render the hallucination a visual hallucination of *o*. We can see why by examining one of Johnston's own examples (borrowed here for my own purposes). He asks us to consider a scenario in which, while dreaming, one undergoes a visual experience as of a blowfish and takes it be Bill Clinton.[17] On Johnston's proposal, this is an hallucination of Bill Clinton because there is some uninstantiated visual profile that strikes the subject as being Bill Clinton. To see what is wrong with this position, consider another case. Suppose that I dream that someone is walking towards me. In the dream, I initially take that person to be Bill Clinton but as he approaches I realize that he is, in fact, my father. In such a case, I initially *misidentify* the object of which I had an apparent visual experience. There is no temptation to say I was dreaming of Clinton, or that my visual experience was of Clinton, as either of these views would fail to do justice to the intuition

---

will strike the hallucinator as being of *o*, just as would likely (but not necessarily) be the case for an instantiated visual profile experienced while veridically perceiving *o*.

[17] Johnston (2004: 232). There is an interesting question as to whether dreams are akin to visual hallucinations (*cf.* Macpherson (2013), Móró (2010)). For present purposes, I will assume with Johnston that these phenomena are similar enough to be discussed together in this context (see also Macpherson (Ibid), Casey (1976) and Freud (1959) for support of this position).

that I misidentified my father when I took him to be Clinton.[18] Nevertheless, this is precisely what Johnston's first proposed anchoring relation delivers.

In fact, Johnston's view must say that in the beginning of my dream I hallucinated Clinton, but by the end I was somehow hallucinating my father.[19] While such transitions may be possible in hallucinatory contexts, this would not be a faithful description of my experience in *this* case, and would do nothing to explain why my attitudes progressed as they did. A case in which I start out hallucinating Bill Clinton and subsequently have that very object be hallucinated as my father is a case in which my hallucination of Clinton *turns into* my father. By contrast, the counterexample case described above is one in which the object of the hallucination was my father all along. The crucial point is that Johnston's first anchoring relation effectively turns misidentifications of apparent objects of perception into *other* apparent objects of perception, and this is untenable. No anchoring relation should render errors of identification of apparent objects a conceptual impossibility.[20]

---

[18] Thanks to [name omitted for blind review] for drawing my attention to this kind of example.

[19] This is not to say that Johnston's overall proposal cannot make room for some sense in which misidentifications of apparent objects of perception are possible. This is only a consequence of the first anchoring relation considered alone. Johnston himself recognizes that misidentification errors are possible (p. 133 and n. 15) when the proposed anchoring relations are in tension with one another, as would be the case if, e.g. an uninstantiated visual profile caused by a prior thought/perception of *o*, struck the hallucinating subject as being *s*. Unfortunately, without further resources this puts the proposal in the undesirable position of having no principled means for determining whether the hallucination in such cases is of *o* or *s,* or both, or neither. In addition to the considerations I offer in the text, I think this consequence also warrants pursuit of a more substantive proposal.

[20] An anonymous referee observed the following complication. Suppose that *S's* dream ended before she came to recognize the 'misidentification'. We might wonder whether undiscovered identities would infect *every* object-involving dream because, in principle, any dream might have lasted longer and included a phase in which it is revealed that the objects of the dream are not the ones we take them to be. This appears to imply radical skepticism about our knowledge of the object-involving content of our dreams.

There are a few things to say in response to this observation. First, radical skepticism about the objects of our dreams is not obviously inevitable. More must be said about both the individuation conditions, and content determination conditions of dreams. This is an interesting and difficult issue that I cannot properly pursue here but I will mention two options. First, one could hold that dreams have their duration essentially, and so a dream that has less or greater duration than the actual dream would simply be a different dream. In the alternative case envisioned, one would maintain that there simply was no misidentification at all, and so no threat of skepticism. A second (and I think preferable) option would be to

This shortcoming is perhaps what led Johnston to ignore the most plausible characterization of the subject's hallucinatory experience in his own blowfish case. In his case, the subject underwent a visual experience as of a blowfish, just like Macbeth underwent a visual experience as of a dagger. Her visual hallucination was arguably of no *particular* blowfish at all. She made a further error in misapplying an individual concept of Bill Clinton to her experience as of a blowfish; it does not follow that she underwent a visual hallucination of Clinton. In contrast, when Macbeth hallucinated Banquo he made no further identificatory errors. He had an experience as of Banquo and developed attitudes accordingly.

The problem with Johnston's first anchoring relation is that it places inadequate constraint on what it is for there to be a particular such that it strikes the subject that she is experiencing it. And so any visual profile could strike the subject as any particular worldly object, if she is deluded enough. As a result, the anchoring relation collapses object-involving visual hallucinations with other delusory mental states that involve misidentification. Consequently, satisfaction of this condition fails to guarantee that a

---

maintain that, like memory content, (some) dream content is determined by informational chains running back to prior experience. Which objects one's dreams are about will depend on the informational connections to prior experiences, and while it may not always be clear to the dreamer whether or not she is mistaken about the objects of her dreams, this would present no greater skeptical challenge than we face in other memory contexts. Finally, I'll offer a comment on the dialectical situation. I have been arguing that Johnston's proposed anchoring relation has the unacceptable consequence that misidentifications of objects in dreams are impossible. While of considerable independent interest, the referee's observation does not clearly help Johnston's proposal on this point. It is difficult to see how mistakes that are possible while waking (i.e. misidentifying the object-involving content of one's experience) should become impossible while asleep. Given this, (assuming skepticism were to somehow be inevitable) it isn't obvious that skepticism about the object-involving contents of one's dreams is somehow less plausible than an alternative infallibility about the object-involving contents of one's dreams.

particular hallucinatory experience is of the sort in which we are interested here.[21] Next, I turn to Johnston's remaining anchoring relations (3.b.i-ii).

The second and third anchoring relations are both developed in terms of causal connections. According to the first, an hallucination will be a hallucination of an object *o* if the uninstantiated visual profile is caused by a prior perception of *o.* According to the second, it will be an hallucination of *o* if it is caused by a thought of *o* to the effect that it is such and so. Then Macbeth's experience is an hallucination of Banquo in virtue of either his prior visual perception(s) of, or his thought(s) about, Banquo.

While some causal connection may be necessary for object-involving hallucinations, the challenge for Johnston's proposal is to show that an appropriately specified causal connection is a sufficient condition and it is difficult to see how this might be done without some further theoretical resources. Many hallucinations may be possible only in virtue of having causal connections to other mental states and external objects. Johnston's proposal must somehow distinguish the cases like Macbeth's hallucination of Banquo from object-independent hallucinations in some informative way.

Suppose we understand causal connections in terms of counterfactual dependence along the lines of Lewis (1973). Then Johnston's causal anchoring conditions could be developed as follows: S's hallucination *h* is an hallucination of *o* if, S would not be aware of visual profile *p* at *t* if she had not perceived (or thought of) *o* at some earlier time *t'*. The view would then say that Macbeth had am hallucination of Banquo because he would not

[21] Johnston himself recognizes a tension in his proposal and hedges by admitting that there are plausibly some qualitative constraints on hallucinations of particulars (p. 132). However, nothing on these lines will save the account, given that a hallucinator may come to appreciate that she has misidentified an object of hallucination even though it has undergone no qualitative change.

have been aware of the uninstantiated visual profile *p* at *t* if he had not earlier seen or at least had thoughts of Banquo.

There are two serious problems with this condition. First, it fails to distinguish hallucinations of external world particulars from other complex visual hallucinations because it is met by object-involving and object-independent hallucinations alike. Take Macbeth's hallucination of a dagger. One possibility is that his experience as of a dagger was the causal result of having earlier perceived some particular dagger or daggers. In such a case, the causal condition would be met. It does not, however, follow from this that Macbeth was hallucinating a particular dagger or daggers. Had a different (but perhaps similar) dagger played the same causal role, nothing in the character of the hallucination would have been different. Hence, which particular dagger was involved was inessential. In short, the causal condition would be met if Macbeth had simply acquired the ability to represent the general property of being a dagger on the basis of his perception of a particular object or objects, and then subsequently had an hallucination involving that representation, but did not hallucinate any particular dagger.

The second serious problem is that, like Johnston's first anchoring relation, this causal condition also fails to distinguish object-involving hallucinations like Macbeth's hallucination of Banquo from other delusory mental states involving misidentification. Consider again Johnston's blowfish case. It may be that the subject had the hallucinatory experience she did as a causal result of prior perceptions/thoughts of both a particular blowfish and of Bill Clinton. Nevertheless, it would be rather implausible to maintain that, because of this, her hallucination was both a visual hallucination of a particular blowfish and, simultaneously, Clinton. It would at least be less plausible than the alternative I

suggested above. Thus, as formulated, all three anchoring relations fail to guarantee that an hallucination has an external world particular as its object.[22]

Let me summarize the argument in this section. Johnston's representative proposal provides three candidate anchoring relations, all of which fail to guarantee that some object is involved in the subject's hallucinatory state. The first clause places inadequate constraints on the notion of a visual profile striking a subject as a particular object and so collapses object-involving hallucinations with other kinds of delusory mental states that involve identificatory errors. The second and third causal clauses fail to distinguish object-involving hallucinations from other complex (object-independent) hallucinatory experiences and fail to distinguish them from other delusory mental states as well.

These problems generalize. It is not enough that memory has provided a subject with a singular or object-directed element. Nor is it enough that, as a result, there is a particular in the world that the subject takes to be the object of experience. The general line of thinking about hallucinations of particulars has been inadequately developed and I will offer a new proposal in the next section.

4.4 Hallucinating (and Imagining) as Perceptual Remembering

In this section, I aim to sketch a more substantive account of object-involving hallucination. I begin by motivating and specifying the view that such hallucinations are a

[22] Two things are worth mentioning here. First, these problems would not be avoided with a more refined counterfactual account of causal connections (*cf.* Lewis, 2000) for requiring a chain of stepwise influence from an experience or thought of *o* to the awareness of an uninstantiated visual profile *p* would do no better than the simple counterfactual account in distinguishing the hallucinations of interest from the others states discussed here. Second, one might worry that, while seeing *o* is an extensional relation, visually representing *o* is not, and so it is really plausible that this is simultaneously an experience of a particular blowfish and Clinton because the subject takes the blowfish to be identical to Clinton. Thanks to [name omitted for blind review] for pressing this issue. I'm not convinced. Suppose that you mistakenly think that the George Harrison that designed pipe organs was also the Beatle. Now suppose, you call to mind a visual image of Harrison on the cover of the "Let it Be" album. You have visually represented the Beatle but it is not clear that you have also visually represented the organ designer.

species of perceptual remembering. I then show how this proposal provides resources for avoiding the problems raised in section one and discuss two further payoffs. As the claim that certain hallucinations are themselves a species of remembering may strike some as counterintuitive, I defend the thesis against objections in section three.

The general line on object-involving hallucinations is on the right track in maintaining that such hallucinations depend on other mental states. It stumbles because it does not do enough to characterize the nature of that dependence. I will argue that we have good reason to accept the claim that object-involving hallucinations are an unusual species of perceptual remembering.[23] We can see why by carefully considering exactly what object-involving hallucinations do for the hallucinating subject.

The shared starting point is that object-involving hallucinations relate the hallucinating subject to external world particulars. However, the discussion so far has revealed that not just any relation will do. When Macbeth hallucinated Banquo, it was not simply because he stood in the 'thinking of' relation to Banquo while hallucinating, for one can deploy individual concepts of particulars while hallucinating without thereby visually hallucinating those particulars (*cf*. the blowfish case). Nor did he hallucinate Banquo by merely having his present hallucination be caused by prior experiences or thoughts of Banquo, for antecedent experiences and thoughts may well have caused Macbeth to undergo a variety of hallucinatory experiences, many of which would not have been object-involving at all (*cf*. the dagger case).

[23]There is some precedent for something like this suggestion in the psychological literature. For example, Penfield (Penfield and Perot (1963)) suggested that *all* hallucinations are reactivated memories. This is implausible (*cf.* Collerton *et al*. (2005)). However, if I'm right, Penfield was not entirely wrong, he simply overextended his suggestion. The proposal I make here is supported by neurophysiological evidence as well: see e.g. Collerton *et al*. (Ibid), Ingle (2005), Mast (2005), Squire (2004), Noda *et al*. (1993), Levine and Finklestein (1982), Sacks (1970) and especially Vignal *et. al.* (2007) who make the case for concluding that certain hallucinatory states at least have the same content as certain episodic memories.

In hallucinating Banquo, Macbeth was related to that particular man in a distinctively *visual* way; he had a *visual experience* as of him. And although this relation was caused by some antecedent experience(s), this is not enough. Its status as an hallucination in which Banquo is the object, must depend on something else as well. The general approach to object-involving hallucination failed to capture the fact that the relationship between hallucinations and prior experiences, whatever the causal relation amounts to, preserves the channel of information between the source (Banquo) and the present experience; that is, a visual hallucination of Banquo necessarily preserves Macbeth's prior perceptual connections to him.

While object-involving hallucinations may be an unusual means of preserving perceptual connections to particular objects, the realization that it does preserve such connections is crucial because there is a more familiar way to do so; namely, by perceptually remembering those objects.[24] The significance of this parallel should not be overlooked. Consider what would have been the case if Macbeth were to have perceptually remembered Banquo. He would have brought to mind an image that, in some sense, matched his prior perceptual experience(s) of that particular man. And, at the very least, his so remembering would have preserved an intimate connection to his prior experiences of Banquo, and *ipso facto* to Banquo himself.[25] Moreover, on this basis, Macbeth could (and probably would) have also known, or continue to have known, a number of things about Banquo, e.g. that he had such and such color hair, that he was roughly such and such height,

[24]See, e.g. Burge (1993, 1997, 2003), Byrne (2010), Campbell (2002), and Martin (2001) for some further discussion of remembering as a means of preserving perceptual connections to objects.

[25] According to Martin (2001: 267) the relation at issue is one that preserves apprehension of events involving those objects; according to Byrne (2010) and Campbell (2002), it preserves "cognitive contact" with external particulars. While there are significant challenges to developing such proposals, I need take no stand on the particular nature of this intimate relationship at this juncture.

and so forth. *Perceptual remembering is a vehicle for the transmission of objectual (and possibly propositional) knowledge of past objects of experience*.

Now return to the case of interest—Macbeth's hallucination of Banquo. In so hallucinating, Macbeth had a visual experience involving an image that, in some sense, matched his prior perceptual experience(s) of Banquo. Although I have so far suggested that such hallucinations preserve perceptual connections, it is worth noting just how intimate the preserved connection is. Macbeth was connected to Banquo via his prior perceptions in such an intimate way so as to be capable of taking Banquo to be present in the scene before his eyes. Moreover, on the basis of his hallucination, Macbeth was able to know (or perhaps continued to know) a variety of things about his former ally (e.g. his hair color, height, and so on). Consequently, like perceptual remembering, *object-involving hallucination is a vehicle for the transmission of objectual (and possibly propositional) knowledge of past objects of perception.*[26]

Finally, suppose that the issue of whether Macbeth remembered Banquo was somehow brought to his attention. He could, without any error, report himself as remembering Banquo in both of the cases described above. In the hallucinatory case, the considerations he would take to justify his memory claim are ostensibly the very same ones

[26] An anonymous referee raised the following objection. If Macbeth takes himself to know facts about Banquo on the basis of perception, but in fact does not do so but rather believes the facts on the basis of remembering, doesn't this undermine his knowledge? Isn't the misconstrual of his route to knowledge sufficiently severe so as to threaten the claim to know?

While I have been focused on first-order knowledge, this objection raises interesting issues about the relationship between first order and higher order knowledge. Though I cannot address this topic properly here, it is clear that subjects' confusion in such cases would undermine some of their self-knowledge. It is not obvious, however, that such errors would impact first-order knowledge as well. One's take on this may depend on one's background epistemological commitments and those endorsing certain externalist theories of justification could arguably deny that failures of higher order knowledge (of this sort) would infect one's first order knowledge. Whether one knows that *p* depends largely on e.g. the reliability of the process by which one's belief that *p* is formed. What one believes about how one came to any particular belief is simply another matter.

he would use were he to remember Banquo *upon actually seeing him*. Collectively, these considerations show that such hallucinations connect the hallucinating subject to prior objects of perception in much the same way that perceptually remembering those objects does. Specifically, both are vehicles for the transmission of knowledge of past objects of perception. Consequently, we have good reason to accept the claim that object-involving hallucinations are a species of perceptual remembering.

Of course, if we are to move beyond the inadequate general line of thinking about object-involving hallucinations by understanding them to be a species of perceptual remembering, more must be said as they are far from prototypical cases of remembering and no substantive characterization of object-involving hallucination should lose track of its uniqueness. The following is a basic proposal:

> A subject *S* visually hallucinates an external world particular *o* just in case *S* perceptually remembers *o* and, instead of presenting as of something past, that memory presents itself to *S* as an occurrent visual perception of *o*.

Though simple, and surely in need of refinement, the proposal has many virtues. First, it makes sense of Macbeth's hallucination of Banquo for on this view it is simply an instance of his perceptually remembering the man. Of course, it is also a pathological kind of remembering because instead of being voluntary and presenting itself as a relation to a past object of perception, it is involuntary and presents itself such that Macbeth takes Banquo to be present in the scene before his eyes.

Second, the proposal is not subject to the objections presented in section one. Experiences like Macbeth's hallucination of a dagger plausibly depend on memory processes in crucial ways. However, they are not instances of remembering particular

objects and they are not visual experiences as of any particular external-world objects. Since the proposal under consideration characterizes object-involving hallucination as a kind of perceptual remembering, it distinguishes the phenomenon from other kinds of hallucinatory experience.

Moreover, being able to refer to an object and being able to deploy a concept of that object both fail to entail perceptually remembering it. For example, I can refer to Cicero and I plausibly have a Cicero-concept available for me to deploy. I cannot, however, visually hallucinate or perceptually remember him for I cannot even visually represent him—I have never seen him, any visual depictions of him, or any detailed descriptions of his physical characteristics. The kind of relation I stand in to Cicero does not preserve a visual information channel between him and my current experience; it does not preserve a perceptual connection to him. So, on the proposal considered here, not just any case of hallucination involving a reference to, or a concept of, a particular will be a visual hallucination of that particular. This allows the proposal to distinguish between object-involving hallucinations and other hallucinatory/delusory states that involve identificatory errors.

To further illustrate, consider again Johnston's original blowfish case in which a subject has an hallucination as of a blowfish and takes it to be Clinton. My proposal predicts (correctly I think) that Clinton is not the object of the subject's visual hallucination. Given how the case is described, there is no temptation to think that the deployment of a Clinton concept preserves a visual connection between the hallucinating subject and Clinton. Consequently, she is not undergoing a perceptual memory (unusual or otherwise) of Clinton. Similarly, if while dreaming, I undergo an experience as of someone walking

toward me and mistakenly take it to be Clinton, before realizing that it is my father, if I hallucinate anyone, I hallucinate my father; I do not hallucinate Clinton.

So far, we have seen how the view that object-involving hallucinations are a species of perceptual remembering allows us to put a principled and more satisfying constraint on what it takes for an external world particular to be an object of hallucination. In doing so, it advances our understanding of the phenomenon by drawing important distinctions where prior treatments of it did not. This allows it to escape the objections raised in section one and such results would be enough to justify the proposal's position as a valuable contribution to our knowledge of hallucination. However, the view has other interesting payoffs as well. I will discuss two more in this section.

First, understanding object-involving hallucinations to be unusual cases of perceptual remembering positions us to provide a philosophical account of a variety empirically observed phenomena. For example, cortical stimulation and spontaneous epileptic seizures can facilitate the repeated visual hallucination of particular scenes (sometimes pleasant, sometimes frightening) from subjects' childhoods. The proposal developed here supports the empirical suggestion that such experiences are instances of *recalled memories* (*cf*. Vignal *et al.* 2007: 88, 90). Similarly, auditory hallucinations of *known* voices are sometimes experienced in the transition from waking to sleep or vice versa (*cf*. Jones *et al*. 2010). In the framework proposed here, we can understand such cases to be unusual cases of *remembering* someone's voice.[27] Finally, in a pathological context, the proposal allows us to understand some patients' phantom limb experiences, PTSD flashbacks, and autoscopic hallucinations of the self (*cf*. Sachs (2012, 1970), APA 1994, and

[27] See Hill and Linden (2013) and Jones *et al*. (2003) for discussions of related auditory hallucinations in non-clinical populations.

Penfield (1963, 1959)) to be unusual memories of one's body (parts), and prior experiences (including of the self) respectively.[28]

Dennett (1991) and Blume *et al*. (1993) point out that the relative veracity of these phenomena presents a challenge to understanding them in any satisfactory way. We might wonder, for example, to what degree an hallucinated object, event, voice, or quality can differ from the source object and still remain an experience as of *it*. Understanding hallucinations of particulars to be instances of an unusual species of perceptual remembering provides an *in principle* way to address this problem: the constraints on these hallucinations should be modeled on the constraints for perceptual remembering, whatever they turn out to be.[29]

Second, the proposal made here can also be extended to understand a variety of other puzzling mental phenomena. In its most generic form, the proposal makes sense of a certain subset of hallucinations by claiming that the hallucinating subject is in a familiar kind of mental state (perceptual remembering) and the delusion consists in her misidentifying the state (as an instance of perceiving).

This is a familiar strategy.[30] For example, Currie (2000) has argued that hallucinations (in general) are just cases in which subjects have mistaken an *imaginative state* for a perceptual state. The immediate difficulty with such a proposal should be

[28] It is worth noting explicitly that there is precedent for treating PTSD flashbacks as a species of remembering in the psychological literature (*cf.* Carroll and Carroll (2005)).

[29] So far, I have suggested that preservation of a perceptual information channel between objects and current mental states is one necessary condition on perceptual remembering. This issue certainly deserves further treatment but that is a project best suited for a general discussion of perceptual memory and space precludes any proper attempt here. See, e.g. Fish (2009), Bernecker (2010), Michaelian (2011), Martin (2001), and Martin and Deutscher (1966) for some preliminary discussion of this issue.

[30] See e.g. Currie (2000: 197-181), Sass ((1994: 19) citing Howard Searles's earlier research), Bentall (1990) who proposes that hallucinations are the product of mistaking an internally generated representation for one received from an external reality, and Morrison (2001) and Morrison *et al.* (2002, 2003) who hypothesize that the problem is related to source monitoring impairments in the episodic memory system.

transparent. Some imaginative states are object-involving; given this, the challenges raised in section one reemerge. For example, we can ask what it would be for Macbeth to imagine Banquo standing beside him, and how that kind of object-involving imaginative state would be distinguished from object-independent imaginative states (e.g. Macbeth's imagining a dagger floating in front of him).

Just as with object-involving hallucination, deploying an individual concept in connection with imagining, or having one's imaginative state depend causally on prior experiences/thoughts of particulars will not suffice for making an imaginative state object-involving. One may only be able to imagine a generic dagger in virtue of having seen some or other particular dagger before, and one may try (but fail) to visually imagine Clinton by calling to mind an image of Al Gore (or perhaps more radically, an image of a blowfish).

While space precludes a fully general discussion, the proposal I have made can be extended as a considerable improvement over prior proposals. Like object-involving hallucination, object-involving imagination preserves a very intimate sensory connection between subjects and external world particulars and can be understood to be another unusual kind of perceptual remembering.[31]

To summarize, the thesis that object-involving hallucinations are an unusual species of perceptual remembering captures the data that challenged Johnston's first pass at an account, and so stands it as an improvement over the general line on hallucinating particulars. It also provides a philosophical framework for understanding a variety of empirically observed phenomena. Finally, it can be naturally extended to help us understand other puzzling mental phenomena as well (e.g. object-involving

[31]See e.g. Martin and Deutscher (1966: 167) for this diagnosis of an analogous case involving a painter and a scene from his childhood. I discuss this case in more detail in section three.

imagination). [32] Despite these considerable payoffs, the idea that object-involving hallucinations do not merely depend on memory processes but, rather, are actually a species of perceptual remembering may strike some as counterintuitive and I turn next to defending the proposal.

4.5 Defense of the Remembering Account of Hallucination

The proposal sketched above reframes the challenge of giving an account of object-involving hallucinations as a special challenge for an account of perceptual remembering. Understanding these hallucinations in this way provides a firmer ground upon which to build and is a substantial improvement over prior approaches to the phenomenon. Nevertheless, one might be tempted to object to my proposal by arguing that object-involving hallucinations lack some essential characteristic or characteristics of memory and thereby cannot be a species of remembering. I will consider three versions of this objection: (i) an 'epistemic transparency' version that focuses on the claim that distinctive phenomenological markers make the fact that a subject is remembering transparent to her; (ii) a 'connection with reality' version that focuses on the claim that hallucination puts subjects "out of touch" with reality; and (iii) an 'autobiography' version that focuses on the claim that one can only remember objects and events they have witnessed.

All three versions attempt to demonstrate that perceptual remembering and object-involving hallucination differ in some essential feature and my strategy for responding to each of them will be the same. I will argue that there are uncontroversial cases of

[32] To stave off some potential confusion, committing to the claim that object-involving hallucinations and object-involving imaginative states are both species of perceptual remembering does not commit one to the further claim that there are no differences between imagining, hallucinating, and *other* species of perceptual remembering (e.g. episodic memory). One obvious difference is that hallucination is the only one that need involve some kind of delusion.

remembering that lack the feature(s) my opponents claim are both essential to memory and absent in hallucination. This shows that perceptual remembering and object-involving hallucination *share* their most crucial features. In doing so, it supports the thesis that object-involving hallucinations are an unusual species of perceptual remembering.

Epistemic Transparency

Remembering perceived objects, events, or qualities is typically thought to involve an awareness of them *as in the past*.[33] One might even suppose that such awareness is the *hallmark* of remembering; it makes the fact that a subject remembers transparent to her. Since hallucinations lack such awareness—when a subject hallucinates, it seems to her as though she *perceives* something and the apparent awareness involved in perception concerns things *in the present*—an opponent might be tempted to conclude that the proposal must be mistaken; no species of hallucination could be a species of remembering.

One could find inspiration for such a position in, e.g. Martin (2001: 271-281). There he argues that phenomenological differences between perception, imagination, and memory are best understood in representational terms. The phenomenology of perception arises from its presentational nature; it *presents* objects, events, and qualities to the mind. Consequently, for subjects undergoing a perceptual experience, it is as if the objects of the experience *must exist and be present*. In contrast, the phenomenology of memory and imagination arises from their *representational* nature; they represent *experiences* of objects, events, and qualities, and this allows those objects, events, and qualities to be

[33] Tulving (1972) introduced the slightly problematic metaphor of "mental time travel" to capture this aspect of remembering. However, the idea that some remembering involves an awareness of objects of memory as in the past is much older—see, e.g. Russell (1921: chapters 9 and 12), James (1890), Locke (1690), and Aristotle's *De Memoria et Reminiscentia* See also Sutton (2004/2010), Bernecker (2010: 13, 14), Squire and Kandel (1999: 106); Brewer (1996); Tulving (2002: 2); Hoerl (1999: 235); Campbell (1994, 1997); Martin (2001) and Schacter *et al*. (2000: 628) for further discussion.

before the mind in a way that does not require that it is as if they must exist and be present. In visually imagining, a subject represents objects, events, or qualities, *as they would be experienced*. In remembering, a subject represents specific prior experiences and in doing so represent the objects that were proper parts of the contents of those experiences. That is, subjects recall previously experienced objects, events, and qualities, *as having been once presented to their point of view.*

As formulated, Martin's proposal is a thesis about how to understand phenomenological differences between these states, *when there are such differences*.[34] However, one might be tempted by a stronger view. Namely, one might hold that such phenomenological differences are *essential* to these respective mental state types.[35] Consequently, if a subject perceives *o*, the distinctive phenomenology of perception enables her to recognize that she perceives; if she remembers *o*, the distinctive phenomenology of memory enables her to recognize that she remembers; and if she imagines *o*, the distinctive phenomenology of imagination enables her to recognize that she imagines.[36] This view would be inconsistent with the proposal I advocate in this paper and, while I can see its initial pull, I think that there are decisive reasons against it. The central problem is that our capacity for self-knowledge of this sort is imperfect and so does not satisfy the demands of such an ambitious proposal.

---

[34] Martin (personal correspondence, and (2001:261))

[35]Although Martin himself does not defend this kind of view, I thank an anonymous reviewer for drawing my attention to the role his discussion could play in motivating such a view.

[36] Martin (2001: 281) does accept that the phenomenological features described above play this role, when present. It is worth noting that, though the phenomenology of memory and its epistemic transparency are generally treated together, they are logically distinct and one could maintain that there are *two* objections here (I thank [name omitted to preserve blind review] for pressing this issue). This possibility will not matter here, as my responses address both dimensions of the issue.

Consider perceptual remembering. We cannot accept a view according to which *all* instances of perceptual remembering are epistemically transparent because there are clearly cases in which subjects remember without recognizing that they are remembering. For example:

> Suppose that someone asks a painter to paint an imaginary scene. The painter agrees to do this and, taking himself to be painting some purely imaginary scene, paints a detailed picture of a farmyard including a certain colored and shaped house, various people with detailed features, particular items of clothing and so on. His parents then recognize the picture as a very accurate representation of a scene [that] the painter saw just once in his childhood…the painter did his work by no mere accident. (Martin and Deutscher (1966:167-8)

Although the painter sincerely believes that his work is purely imaginary, there is good reason to think that he has actually remembered, and then painted, a scene from his childhood.[37] As with other instances of perceptual remembering, the painter's mental states are a vehicle for knowledge of past objects of perception. In this case, his knowledge of the past scene is simply expressed by his painting.[38]

Other cases of non-transparent perceptual remembering are not particularly hard to come by. To see one further example, consider the familiar scenario in which you are unsure whether you are remembering a prior event, or whether you are simply imagining that some event had previously occurred. In such cases, you are not aware of the event as being in the past and you are not aware that you are remembering (both are precisely what you are unsure of) and you lack this awareness even if you are, in fact, remembering.[39] So, perceptually remembering something need not be epistemically transparent; one can

[37] Most accept this assessment of the case; see, e.g. Deutscher (1988: 58), Byrne (2010), and Sutton (2010/2004). There are exceptions, e.g. Debus (2010).

[38] There is, of course, a sense in which this expression of knowledge is peculiar. This arises from the fact that, though the painter clearly had knowledge of a past visual scene (such knowledge is necessary for his painting as he did), he did not obviously have doxastic propositional knowledge about the scene. Among other things, he failed to know (or remember) *that* there *was* a certain colored and shaped house in a farmyard.

[39] See Bernecker (2010: 88-90) for an analogous case involving remembering *facts* about prior events. See also Matthen (2010), and Russell and Hanna (2012) for discussion of other cases.

remember without recognizing (or without it seeming to one) that one is remembering. Consequently, the fact that subjects lack awareness of objects *as in the past* in cases of object-involving hallucination cannot be used to argue that those hallucinations must not be a kind of perceptual remembering.[40]

Examples in which phenomenology fails to lead to appropriate metarepresentational judgments are not special to memory. For example, subjects are capable of judging that perceptual experiences are familiar from the past when they are not (as in déjà vu) or vice versa (as in, e.g. failures to recognize known faces).[41] They are also capable of mistakenly judging perceptual experiences to be instances of perceptual imagination. [42] This suggests a more general problem for views that make phenomenological features essential to the criteria by which to draw metaphysical distinctions between mental state types (at least in the cases of perception, imagination, and memory). Namely, such views assign the wrong theoretical role to phenomenology.

It is not surprising that we can, in various circumstances, be mistaken about whether or not we are perceiving, imagining, or remembering something. Psychological processes are generally imperfect and there is little reason to think that the processes

[40] One might try to get at the same objection by claiming that while all perceptual remembering is of things that are not present, object-involving hallucinations are hallucinations of the *presence* of some entity. From this, my opponent might be tempted to conclude that hallucinations cannot be instances of perceptual remembering. (I thank an anonymous reviewer for pushing this line of objection). The problem with this formulation of the objection is that the hallucination of the presence of an entity doesn't entail the presence of an entity, at least on most treatments of hallucination. So it is perfectly plausible that while one hallucinates the presence of an entity when undergoing an object-involving hallucination, her hallucination is still *of* something that is not present, just as are cases of perceptual remembering.

[41] Failures of recognition can come in several varieties. Some are relatively mundane, as when you might simply fail to recognize someone on a particular occasion. Others are more pathological as when, e.g. some subjects suffering from Capgras Syndrome seem to, in some sense, recognize familiar individuals but see them *as unfamiliar* (see, e.g. Davies and Coltheart (2000) for some discussion).

[42] This was arguably demonstrated in a research setting by Perky (1910). That subjects can make such mistakes does not necessarily entail that, in such cases, there is no phenomenological difference between perceiving and imagining (see Hopkins (2012) for some discussion). However, it does demonstrate that perceiving does not entail realizing that one perceives.

underlying the phenomenological features that guide meta-representational judgments about mental states should somehow be different. Given this, phenomenology is not best seen as a *determining* factor for mental state types. Rather, a more plausible theoretical role for phenomenology is as *defeasible evidence* of the type of mental state occupied by a subject on a given occasion. While the fact that subjects are sometimes led astray by phenomenology challenges its potential role as a component of mental state type definitions, it does nothing to challenge the idea that, when subjects do recognize that they are, e.g. remembering, it is typically because it seems to them that the objects of their experience are in the past. One simply should not infer from this that all perceptual remembering involves an awareness of objects as in the past. Consequently, one should not conclude that object-involving hallucination is not an unusual species of perceptual remembering on phenomenological grounds.

Connection with Reality

A second version of the objection focuses on a different apparent epistemic discrepancy between hallucination and memory; namely, an opponent might claim that hallucinating necessarily puts one "out of touch" with reality, whereas remembering puts one, in some sense, in touch with reality. Consequently, she might say, object-involving hallucinations cannot be instances of perceptual remembering.[43]

This objection seems to rests on an overly restricted conception of hallucination. Object-involving hallucinations *do* put subjects in touch with parts of reality—that is what makes them a distinctive species of hallucination. They connect subjects to reality by preserving a perceptual connection to real things (past objects of perception). They are

[43] I thank an anonymous reviewer for pushing this objection.

hallucinations because, in addition to the knowledge they sustain, they also present unreal states of affairs and in particular, they mis-present the *temporal source* of represented objects, events, or qualities by presenting them as present when they are not.

For the objection to put pressure on my proposal it would have to be the case that perceptual remembering never misrepresents the temporal source of the objects of experience. However, this is clearly not the case. It is not only possible for memory to provide mistaken information about the temporal source of objects of remembering, but it is likely that such errors occur relatively frequently.

To illustrate, suppose that three years ago, you attended all three APA division meetings and went to numerous talks at each of those meetings. Importantly, you attended a talk by a particularly famous philosopher at the Pacific. Now, three years later, suppose that you recollect the famous philosopher's talk but mistakenly take it to have occurred at the Eastern. In such a case, there is an object you remember (the philosopher's talk), despite being out of touch with a particular feature of reality; namely, despite being wrong about the temporal source of the object represented by your memory. Perceptual remembering tolerates error in ways that propositional remembering does not.[44] Consequently, we cannot conclude that object-involving hallucinations are not instances of perceptual remembering because they put the subject out of touch with a certain feature of reality.

One might try to press this challenge further by arguing that the problem is not just that one gets the temporal source of the represented objects wrong when hallucinating.

[44]For example, while this kind of error provides no challenge to remembering the talk, it does preclude you from remembering *that* the talk was at the Eastern.

Rather, the problem is that hallucination involves a particular kind of error—the object presents as though *from the present*.

This consideration would only challenge my proposal if it were impossible for a remembering subject to take a remembered object to be present. We have already seen that representing an object *as past* is not necessary for remembering. A little consideration makes it clear that one can also remember while representing an object *as present*. This is exactly what happens when one recognizes someone that they have already seen before. In such cases, the remembered object is taken to be present because it *is* present.

One might note that *recognition* is itself a complicated matter and there is a sense in which what one sees before them must somehow *also* be presenting as past for the subject feels that the object she sees is *familiar*.[45] However, this provides no challenge to the present account because the object-involving hallucinations in which we have been interested *share* this feature with perceptual remembering. In hallucinating Banquo, Macbeth took him to be familiar just as one who sees someone they *recognize* takes that person to be familiar.

So one cannot drive a wedge between object-involving hallucination and perceptual remembering by citing the presence of mistaken information about the temporal source of the remembered object. The upshot is that one cannot deny that object-involving hallucination is an unusual species of perceptual remembering by arguing that hallucination puts one out of touch with reality. Object involving hallucinations are

[45]To avoid a possible misunderstanding, it is worth pointing out that the fact-recognition tasks and free recall tasks can be dissociated is not obviously evidence that recognizing does not involve perceptual remembering. At the very least, recognition seems to be mediated by the processing of perceptual features (e.g. colors and borders) *cf*. Hasselmo (2012: 140). See also Dopkins *et al.* (2013) for some empirical support for the claim that perceptual information is (sometimes) involved with recognition that is based "only" on the distinctive feeling of familiarity.

distinctive because they put subjects in touch with certain parts of reality—they are a vehicle for knowledge of past objects of perception, just as are other species of perceptual remembering. Moreover, those senses in which object-involving hallucinations misrepresent reality are can be shared by perceptual remembering as well.

Autobiography

Finally, a third version of the objection could be made as follows. Perceptual remembering is always the remembering of something in one's past (i.e. it is autobiographical). However, an opponent might argue that one can have an object-involving hallucination of something that is not in one's past. Johnston (2004: 122) offers one putative case of this in the form of a "Manchurian Candidate" scenario. Suppose, for example, that Macbeth had never actually interacted with Banquo before his hallucinatory experience in the dining hall. Instead an evil (time travelling) neuro-surgeon implanted representations of him in Macbeth's mind. In such a scenario, my opponent might be tempted to argue that Macbeth could have still undergone a hallucination of Banquo exactly as he did as described in section one, but he could not have been remembering him in virtue of the fact that he had had no prior experiences of him.

There are at least two responses to this kind of case. First, one might question whether this is actually an hallucination of Banquo as opposed to an object-independent hallucination (i.e. an hallucination of a man but no man in particular) that is phenomenologically indistinguishable from an hallucination of Banquo. In the absence of a fully developed theory of object-involving representation, it is simply not clear whether such implantations can preserve connections to particulars. Second, suppose that we grant that this is genuinely an object-involving hallucination of Banquo. Even granting this, I

think there is good (independent) reason to reject the first premise. Specifically, while most cases of remembering objects/individuals are autobiographical it is not clear that *all* are. We often speak of remembering things with which we have no prior personal experience. For example, we remember Dr. Martin Luther King Jr. and many of us do so in a distinctively visual way—we call to mind images of the man—despite not having prior direct perceptions of him. The fact that television or some other form of visual information transmission mediates our visual representations of Dr. King surely stands as no obstacle to our remembering him.

Given that at least some cases of perceptual remembering are consistent with mediated visual information, we might simply take the fantastical evil-neurosurgeon scenario to be a case in which there is an unusual intermediary in the process of transmitting visual Banquo-information to Macbeth. That is, in so far as we think that such cases are object-involving, it is also plausible such cases involve unusual perceptual memories. As such, we cannot straight away conclude that Macbeth cannot remember Banquo in the scenario sketched here, and this version of the objection, like those before it, does not present a counterexample to my proposal.[46]

The upshot of this section is that the putative counterexamples to my thesis fail because remembering is a multifaceted cognitive ability. Although it often involves a sense

[46] Two further points of clarification are necessary. First, I accept that there are certain Banquo-related things that Macbeth could not remember in a Manchurian Candidate scenario. For example, he could not remember *seeing* Banquo, *going* to battle with Banquo, or so forth, for such events did not happen. In contrast, Banquo did exist as an external particular, and so unlike these fictional events, can be the object of a remembering state (similar points apply to my remembering of Dr. King). Second, one cannot circumvent my response by trying to construct the case as one in which Macbeth is having an object-involving hallucination of a prior *event* that he never experienced. Recall from my earlier discussion that object-involving hallucinations can come in a variety of forms (ranging from partial to full hallucinations, and from partially to fully object-involving). In a case in which a neuroscientist plants a "memory" of a fictional event involving Macbeth and Banquo into Macbeth's mind, my proposal would predict (correctly I think) that Macbeth might subsequently undergo an hallucination of Banquo in the context of a larger hallucinatory experience that involved both object-involving and object-independent event elements.

in which the remembered objects are remembered as past, it can involve a variety of phenomenological features (or, as in the case of the painter, the lack of any distinctive phenomenological feature at all).[47] Moreover, while some or other causal relation between remembering subject and past object/event is surely necessary for remembering, memory can arguably be sustained by a variety of causal relations between subjects and remembered objects. Some of those relations make the connection more direct (*cf*. the painter case), and others make it more mediated (*cf*. our memories of MLK Jr.) Given the heterogeneous nature of perceptual remembering, my opponents' cases fail to challenge the idea object-involving hallucination is simply one of its more unusual species.

The upshot is that this proposal provides a sketch of a satisfactory account of the anchoring relation between one's hallucinatory state and her prior mental states in part because it distinguishes object-involving hallucinations from other kinds of complex visual hallucinations that depend on memory, and from other kinds of delusory states that involve identificatory errors. It also puts us in a position to explain a number of empirically observed phenomena e.g. auditory hallucinations of *known* voices, and provides a model for understanding other puzzling mental phenomena, e.g. object-involving imagination.

To summarize this chapter, I have argued for three claims. First, the preliminary reticence to the view that object and event remembering is amenable to a generative epistemic account is misguided. Second, to remember an event or object is to have object-involving knowing of it by operations of memory. And third, surprisingly, object-involving visual hallucinations and imaginations are amenable to this account as well as they are a species of unusual perceptual remembering.

[47] See Bernecker (2010: chapter 1) for some discussion of the spectra of phenomenology and imagery in memory.

Chapter 5 Does G.E.T. Get the Nature of Memory Wrong?

According to the Generative Epistemic Theory of remembering, memory is a source, and remembering a species, of knowledge. This final chapter responds to four objections.

Having dealt with preliminary opposition to the idea that one can know objects or events by remembering them in chapter 4, I return to fact remembering and focus the remaining discussion on the heart of opposition to G.E.T. – arguments that have the following general structure. (Premise 1) If to remember that *p* is to know that *p*, then every (propositional) output of memory is knowledge. (Premise 2) It is not the case that every (propositional) output of memory is knowledge. (Conclusion) It is not the case that to remember that *p* is to know that *p*.

Premise 2 is uncontroversially true. However, the argument fails because premise one is false. Nevertheless such reasoning appears to have led some to resist epistemic theories of remembering and I explore four ways one might go astray. While they surely do not exhaust the objections one might make, they are representative and their failure to undermine G.E.T. supports the conclusion that the view developed in chapters 3 and 4 *does not* get the nature of memory wrong. Rather, the theory provides a compelling account of the relationships between memory, remembering, and knowledge.

In the first three sections, I respond to *a posteriori* arguments that G.E.T. misidentifies the nature and purpose of human memory faculties, and in doing so requires knowledge where we should not (always) expect it to be. In the fourth section, I respond to an *a priori* argument that G.E.T. is subject to counterexample.

5.1 Memory is not for remembering

Some might be tempted to think that recent empirical work on the psychology of memory shows that G.E.T. misidentifies the nature and purpose of human memory.[1] It is now generally accepted that memory is malleable. Consequently, it provides an imperfect record of the past and acquired facts and its deficiencies arise as a natural consequence of the operation of the processes and mechanisms on which it is based (Moscovitch, 2007: 19-20). Moreover, it is not clear that accuracy or completeness of memory representations is always beneficial, especially from an evolutionary perspective. Accurate and complete recollection can hamper processes of information generalization and categorization and so can actually be a burden for a system that evolved to allow us to function effectively as organisms navigating the world (Dudai, 2002: 95).

Felipe De Brigard has recently used such empirical observations to argue that memory is not best understood as a means for remembering (at least as it is usually conceptualized), but instead as part of a larger cognitive system that supports three things: (i) thinking of what *was* the case; (ii) thinking of what *could be* the case; (iii) thinking of what *could have been* the case (De Brigard, 2014: 158). That is, according to De Brigard, memory and remembering are a matter of *mental simulation*, not necessarily faithful representation of facts about the past or other actual states of affairs.

His central motivation is that many instances of human misremembering, i.e. cases in which subjects fail to recollect things as they actually were, but instead

[1] See Schacter and Addis, 2007, Schacter, 2001, and Schacter *et al*. 1998 for some general discussion of the empirical literature.

misrepresent them in one or more ways, are far from detrimental to those subjects. This, he thinks, challenge theorists to do two things. First, it puts pressure on us to reassess the normative standards by means of which successful recollection is measured. Specifically, we must reconsider whether fidelity to the past requires something less than truth and accuracy of representation (*cf*. Hazlett, 2010; Campbell, 2006; Schechtman, 1994). Second, it pushes us to re-evaluate the function or purpose of human memory, and its relationship to our ordinary conceptions of successful and unsuccessful remembering (De Brigard, 2014: 156-9).

The basic idea is that the empirical study of memory has shown that misremembering is a common phenomenon, easily generated in laboratory settings and frequently occurring in everyday life. Moreover, entertaining false memories is not inherently problematic, but sometimes even beneficial and often the result of normal/healthy brain functioning (De Brigard, 2014:162-4).[2] Given this, De Brigard thinks it would be a mistake to view *all* misremembering as the result of some kind of *malfunction*. In short, memory outputs can be *epistemically defective* in the sense of not meeting an appropriate mind-world relation (e.g. misremembering a yield sign for a stop sign in a remembered scene) without being produced by a memory system that is somehow operating less than optimally (2014:279). This, he thinks, justifies reframing our understanding of the purpose of memory and remembering in terms of their functional roles in a cognitive system designed for *hypothetical reasoning* and event simulation. If this is right, then one might argue that G.E.T.

[2] I discuss some specifics of misremembering in more detail in section 3.2.

must be false because knowledge as such (and in particular knowledge of the past) is not the biological and evolutionary purpose of memory.

The first observation to make is that one cannot infer that accurate recollection is not *a* purpose or function of memory systems from the observation that the systems frequently deliver inaccurate representations. Frequency of a mechanism's success is orthogonal to what counts as a success. The odds of winning a lottery may be 1 in 75,000,000 and any given lottery player will likely never win, but winning is nevertheless *a* purpose of playing the lottery. In baseball, the best batters get fewer than 4 hits out of every 10 at-bats, but getting a hit is still a purpose of batting. Similarly, remembering, *qua* species of knowledge, may be a purpose of memory systems regardless of how frequently one successfully remembers.

One might try to press further by claiming that the problem is not merely that memory frequently produces inaccurate representations without detrimental effects. Rather, it is that once we think about memory (the faculty) in terms of its contributions to organisms' evolutionary fitness, we see that the accurate representation that is a prerequisite for knowledge is not itself a plausible target. Memory systems contribute to an organism's fitness only insofar as they allow that organism to *recast* knowledge of particular events that happened in the past in order to predict what may happen in the future (Suddendorf and Corballis, 2007). And successful anticipatory behavior is correlated with the degree of flexibility the system has to *rearrange* stored information, not necessarily the system's capacity for accurate representation (of the past). Thus, from the evolutionary perspective,

memories (*qua* outputs of the memory system) need not be perfectly faithful renditions of the past – "memory need be only as 'good' as the advantage in decision-making it affords, measured against the cost of its operation" (Boyer, 2009: 514, *cf.* De Brigard, 2014: 165).

Given this, memory is plausibly *designed* (by evolution) to produce incomplete and sometimes relatively inaccurate representations because its psychological utility in doing so is more beneficial to an organism than its correspondence with reality would be (Sutton 2009). And there is something to this thought as inaccurate memories (and beliefs about the past) can sometimes provide emotional, personal, and social advantages to organisms that accurate ones could not (Alea and Bluck, 2003 *cf.* De Brigard 2014:165, and McKay and Dennett, 2009). If the flexibility demanded by the cognitive tasks in which (at least some) memory systems are deployed *requires* relatively imprecise (and sometimes even inaccurate) memory representations, we cannot expect remembering to entail knowing for doing so would require misunderstanding the purpose of memory (De Brigard 2014: 166).

There are a few things to say in response to such an objection. First, there is an important difference between imprecise and inaccurate representations. G.E.T. is not only consistent with the observation that many of our memories are imprecise, but it also predicts as much. Remembering and knowing both come in degrees of specificity and we should expect much of what we remember to be relatively imprecise. For example, one can remember (know) that a car accident involved two cars, while failing to remember (know) that it involved two Fords. That many of our

memories only accurately represent the gist of an event (Schacter and Addis, 2007: 774) is no objection to an epistemic theory of remembering. It shows only that (sometimes) we only know the gist of things by remembering – a point on which all should agree.

The more interesting question is whether G.E.T. is consistent with *inaccurate* representations, and there is good reason to think that it is. Consider a case in which a subject accurately recalls that a car accident involved two cars but also inaccurately seems to recall that the intersection at which they collided had a stop sign (suppose it had a yield sign instead). Such errors likely occur quite frequently (Loftus, 1975) and, on reflection, should not be terribly surprising. The presence or absence of certain traffic signs would not be particularly salient to witnesses of an accident; their attention would naturally be focused on the accident itself. Importantly, it poses no problem for G.E.T. because remembering *what happened* during an event is a complex matter and subjects will be more successful in remembering some things than they are in remembering others. The key is that, in such cases, there is no pressure to judge that a subject *remembers* that there was a stop sign at the intersection; she only seems to remember it. Such an error presents no obstacle to judging that she does remember (and know) a variety of other things about what happened, e.g. that the accident involved two cars. It would simply be a mistake to think that G.E.T. requires memory (the faculty) to be infallible.

A second, and more general, response is that such an objection appears to be a red herring. The prevalence and potential beneficial nature of some misremembering could only be thought to be in tension with G.E.T. if one was also

committed to the thesis that memory systems' *only* subserved remembering (*qua* species of knowledge). Of course, this is rather implausible. Any system can have more than one function, and in so far as memory systems are implemented in neural systems in humans, there is a growing consensus that redeployment of neural systems for a variety of different cognitive tasks is commonplace (*cf*. Anderson 2010, 2007). Consequently, it is plausible that memory systems are put to use for a variety of purposes including hypothetical/counterfactual reasoning (De Brigard, 2014), the construction of personal narratives (Schectman, 1994), otherwise generating a contextually appropriate characterization of the past (Campbell, 2006; Engel, 2000), *and* remembering as characterized by G.E.T.

Far from providing any challenge to G.E.T, observing that memory systems have a plurality of purposes (many of which may require inaccurate representations), reveals some of its theoretical value. G.E.T. provides a principled way to distinguish cases in which one *remembers* that *p* from cases in which one has done something else on the basis of memory. Remembering states are those that qualify as knowledge. Those states that do not are something else, however valuable they may be.

Moreover, one cannot use the observation that inaccurate memory representations can be beneficial to an organism to argue against an epistemic theory of remembering because the same points apply to some cases of failing to know. While epistemologists are often concerned with understanding the value of knowledge, it is also clear that some ignorance (even error) can be practically, socially, and emotionally beneficial to humans. For example, consider a romantic

partnership that lasts for many years. Suppose further that, as with most relationships, the partnership involved various (major and minor) transgressions. It can be in both partners' best interests to simply no longer *know* the details of many of these events. Knowing who said or did what to whom can put undue emotional strain on an otherwise normal relationship.

The general point is that knowledge is valuable, but only to a point. Knowing too much (especially in terms of detail) can be burdensome in a variety of contexts. Consider one further example. Suppose one were trying to plan a trip by car. While knowing which roads have tolls and/or are heavily traveled may be beneficial to one's planning, there are many other things one would do best to not know, e.g. that one particular route would involve 65 right hand turns and 37 left hand turns, or that the trip would span a 9.645599 degree change in latitude and a 22.113578 degree change in longitude, or that the car one will drive is either red or blue, or has either four or three wheels, and so on and so forth). The empirical observations appealed to by this opponent of G.E.T. are a special instance of the observation that humans are finite creatures and excess knowledge (whether remembered or acquired in some other way) can be burdensome for any such creatures.[3]

5.2 Memory connects us to our past *experience and mental states*

Given the arguments just made, one might grant that remembering is a matter of accurate representation but be tempted by the following objection. Remembering is not a matter of knowledge of the world because memory does not connect us to the past itself. Rather, it connects us to our prior experiences and

[3] See Parker *et al.,* 2006, and Lauria, 1987, and Michaelian, 2011 for some discussion in memory contexts, and Harman, 1986 for more general epistemological discussions.

related mental states. Consequently, unlike knowledge, accuracy of remembering is determined only by the accuracy with which it represents a subject's prior mental states.[4]

The central thought is that the information our memory system has to work with is limited by the kinds of input it receives. Consequently, the output of one's memory can fail to constitute knowledge for a variety of reasons. For example, the memory system may have taken a perceptual illusion (Roediger 1996), dream, or fantasy (Loftus, 1996) as input. In such cases, memory itself need have done nothing wrong at all. From the point of view of one's memory, the inputs are real (Dudai, 2002). However, given their epistemically deficient nature, what one remembers cannot be knowledge.

For example, suppose that at t_1 John takes himself to see Nancy leaving the library, and so comes to believe that Nancy left the library at t_1. However, despite the fact that Nancy did in fact leave the library at t_1, John was mistaken in what he saw; he saw *Linda* leaving the library. At t_2, John tries to remember who he saw leaving the library and takes himself to remember that Nancy left the library at t_1. John's memory systems operated flawlessly, in one sense, and produced a true belief about Nancy. Consequently, my opponent might maintain that this is an instance in which John remembers that Nancy left the library at t_1. Of course, it cannot be an instance of knowledge because it does not bear the appropriate relation to Nancy's leaving the library. Nevertheless, my opponent might maintain that it can count as

[4] Thanks to Sven Bernecker and Sarah Robins for this line of objection (personal correspondence).

an instance of remembering because it is faithful to (and caused in the right way by) John's prior belief, and it produced a true memory belief.[5]

It is not clear that such cases should inspire us to think that faithfulness to one's past mental states, and not the past itself, is *the* proper constraint on remembering. First, notice that the very same kind of reasoning could be applied to the formation of true perceptual beliefs. One could, for example, claim that faithfulness to the world is not the appropriate criterion of seeing that *p*. Rather, it is faithfulness to the retinal image, for one's perceptual processes can operate flawlessly and lead one to see that *p*, when one fails to know that *p*.

For example, suppose a subject takes herself to see a man standing off to her left and so believes that there is a man standing off to her left on the basis of her visual experience. Suppose further that there is a man standing off to her left, but that her experience is caused by a sequence of mirrors (of which she is unaware) reflecting a man that is standing off to her right. [6]

Given that the man standing off to her left is occluded by a mirror, the analogous view of perception would hold that since perceiving connects us to our experience of the world, and not the world itself, the subject sees that there is a man standing off to her left, but fails to know that there is a man standing off to her left. Of course, this would be an implausible conclusion to draw. A better assessment is that the subject *seems* to see a man off to her left and so believes that there is a man

[5] See Bernecker (2010: chapter 3) for this judgment about a similar case. There is an interesting question as to whether or not remembering would be a factive state if faithfulness to one's prior mental states is the crucial constraint on remembering. This, however, goes beyond the scope of the present discussion.

[6] See Tye, 2007 and Grice, 1961 for similar cases put to different uses.

off to her left. She does not, however, see *that* there is a man off to her left. Similarly, we might think the most natural assessment of John is that he *seems* to remember that Nancy left the library but does not actually remember that she did.

One might resist by pointing out that the proponent of G.E.T. cannot simply assume that memory and perception are parallel on this matter. While it may be hard to deny that the purpose of perception is to connect us to the world, one might think it is less difficult to deny that memory has the same function. After all, my opponent is *proposing* that memory's function is to connect us to past experiences and mental states. She might maintain that memory *can* connect us to the past, but only indirectly. If one's prior experience put one in touch with the world in the appropriate way, so too might memory (indirectly). If the experience did not, memory will not.

This leads to a second reason to think that this objection is misguided. My opponent is granting G.E.T. that remembering involves accurately representing things. Of course, when one attempts to remember, there are (at least) *two* things that may be at issue: (i) how things were and (ii) how one took things to be. And it is natural for us to distinguish these because we can be interested in either or both when evaluating what one remembers. In short, memory can be a source of knowledge about the world, and it can also be a source of *self-knowledge.*

Consider John's mental state at t_2. Given that he is trying to remember *who left the library*, our judgments about whether he remembers that Nancy left, may be unclear. This is because his current belief stands in no relevant relation to the fact that makes it true. However, notice that if we change the case slightly such that John

tries to remember who *he thought he saw* leave the library, we have no hesitation in judging that John remembers that he thought he saw Nancy leave the library. The natural explanation is that his present belief (about his prior experiences and mental states) *does* stand in a relevant relation to the facts that make his belief true (namely, the fact that he thought he saw Nancy leave). Moreover, the relation is an epistemically sound one. We have no hesitation because when John remembers that he thought he saw Nancy leave, he *knows* something about his prior experiences and beliefs.[7]

Notice that we can get a structurally similar set of judgments about the perceptual case as well. Our judgments about whether the subject sees that there is a man standing off to her left are unclear because her perceptual belief does not bear any relevant relation to the man that is standing off to her left. However, were we to change the case such that the subject formed a belief about how things *seemed* to her to be, we would have no problem reporting that she believed that she seemed to see a man standing off to her left. Again, a natural explanation of this is that her belief about her own experience does stand in a relevant (and epistemically sound) relation to her experience.

The upshot is that both perception and memory can depend on something without that thing being the relevant constraint by which we judge its accuracy (e.g. in perception it may be retinal images; in memory it may be prior mental states and experiences). Consequently, we have not been given reason to think that memory connects us *only* to our prior experiences. It *can* connect us to those prior

[7] One might worry that there is considerable evidence that our memories of our prior experiences and attitudes are not particularly reliable. More on this in section 5.3.

experiences and when it does, we acquire self-knowledge about our prior experiences and mental states (we remember that we held some attitude, or had some experience). It can also connect us to the past itself (via those experiences and mental states) and when it does, we remember, and know, facts about the world itself.

Let me summarize the argument in the first two sections. We saw that the fact that our faculty of memory is flexible, naturally produces misrepresentations, and may be put to use for a variety of purposes is no challenge to an epistemic theory of remembering (Section 5.1). Similarly, the fact that what we remember is (to some degree) constrained by the inputs to the memory system does not show that those inputs provide the (unique) criteria by which remembering must be judged (Section 5.2). Consequently, the issues considered so far give us no reason to think that G.E.T has misidentified the nature and purpose of memory and remembering.

5.3 Remembering as evidence part 1: *a posteriori* arguments

So far, I have addressed the objection that G.E.T. misunderstands the nature of memory itself. Another kind of objection charges it with misidentifying the relationship between remembering and knowing. Specifically, an opponent might be tempted to argue that remembering is not a species of knowledge, but rather it is a species of *evidence*.[8] In this section, I respond to two a posteriori arguments for this thesis.

5.3.1 Reliability

[8] The point is moot if Williamson (2000) is right in thinking that E = K, but I set this possibility aside for the sake of argument.

Since the 1970s, psychologists have uncovered a variety of unexpected (and in some ways worrisome) influences on the accuracy of our memories.[9] The research is so striking that skepticism about the trustworthiness of our memory faculties has become almost fashionable. As one author put it, "As we all know (and science confirms), the act of remembering is not only subjective, but extremely flawed" (Anderson, 2013).[10] Given such a sentiment, an opponent of G.E.T. might be tempted to argue that remembering that *p* cannot be a species of knowledge because our faculty of memory is *unreliable* and so bound to produce instances of remembering that are not also instances of knowledge.

The details of the evidence that memory lacks reliability matter. To paraphrase Descartes, it would be madness to doubt the epistemic virtues of a faculty on the basis of its having deceived us in certain particularly extravagant kinds of cases.[11] An opponent of G.E.T. must identify the sense in which memory is unreliable. She must also show that this lack of reliability undermines the claim that remembering is a species of knowledge, but does not lead to general skepticism about memory knowledge.

An opponent might try to appeal to cases in which subjects make errors (i.e. misremember) more frequently and systematically than we would have expected to support the idea that memory is unreliable. I will consider two representative kinds

[9] See Loftus, 2005 for a review of some of this research.
[10] A Google search of "Memory Reliability" found articles with the following titles (among many other similar ones): "Why we can't trust even our most powerful memories" (MinnPost, 2011), "Trust your memory? Maybe you shouldn't" (CNN, 2013), "Our Memories Are Not As Objective And Reliable as We Think They Are" (*Medical News Today,* 2011), "Memory Not Reliable, Court Says" (*The Scientist,* 2012), "Memory Surprisingly, Unreliable, Study Shows" (*Cosmos the Science of Everything*, 2008)
[11] Descartes, *Meditations on First Philosophy* (Meditation 1; paragraph 3)

of cases, and argue that neither show that memory is unreliable in a way that challenges G.E.T. [12]

Flashbulb memories[13]

Some experiences seem unforgettable, almost as though one stored a snapshot of them as they were happening (e.g. learning of the terrorist attacks on September 11, 2001). Despite the vividness with which subjects seem to remember such experiences, their memories tend to be inaccurate. The problem is that, as we rehearse these experiences in our own minds or with others, our memory processes incorporate additional related information, regardless of whether or not it actually corresponds to our own experiences. Consequently, although we seem to vividly remember where we were and what we were doing, etc. when we first "heard the news" in great detail, we tend to call to mind what *would have made sense to have happened*. This does not necessarily or perhaps even *usually* correspond to what actually did happen.[14] Consider the following case.

---

[12] There are other kinds of cases in which subjects systematically distort memory representations. For example, subjects also tend to "remember" recent events as being more remote than they actually were and remote events as being more recent than they actually were (Neter and Waksberg, 1964); they also "remember" many objects from a wider-angle view than the perspective from which they were perceived, thus making the background within which an object was seen seem disproportionately large (Park, *et al.,* 2007). I set aside such cases because they are more applicable to the remembering of objects and events, as opposed to facts about them. As we saw in chapter 4, knowledge of objects and events plausibly requires different standards than those applicable ordinary propositional knowledge. Moreover, such errors seem consistent with both the remembering and having knowledge of objects and events. One can remember (have knowledge of) an event without remembering (having knowledge of) exactly when it happened and one can remember (have knowledge of) an object without remembering (having knowledge of) the size of the environment in which one saw it.

[13] Brown and Kulik (1977) coined the term "flashbulb memory," though the phenomenon has been recognized in the psychological literature since at least James (1890).

[14] See e.g. Luminet and Curci (2009), Reisberg (2006), Greenberg (2004), and Pezdak (2003) among many others for discussion. It is worth mentioning here that there is no correlation between one's confidence in a flashbulb memory and its accuracy and, in some cases, subject's confidence in their memory experience persists even in the face of counterevidence (Neisser and Harsch 1992)

Former President George W. Bush recounted his experience on 9/11 at least three times between December 1, 2001 and January 15, 2002. Though his recollection differed in detail each time, he consistently recounted seeing footage of the first plane hitting the North Tower *before* he entered Florida schoolroom to read to children. I have included an abridged transcription of one recollection below (*cf.* Greenberg, 2004: 363-4):

At a town hall meeting in December 2001, Bush responded to a child's question "How did you feel when you heard about the terrorist attack?" as follows:

> Well, Jordan, you're not going to believe what state I was in when I heard about the terrorist attack. I was in Florida. And my chief of staff, Andy Card—actually I was in a classroom talking about a reading program that works. And I was sitting outside the classroom waiting to go in, and I saw an airplane hit the tower—the TV was obviously on, and I use[d] to fly myself, and I said, 'There's one terrible pilot.' And I said, 'It must have been a horrible accident.' But I was whisked off there—I didn't have much time to think about it, and I was sitting in the classroom, and Andy Card, my chief who was sitting over here walked in and said, 'A second plane has hit the tower'.

Since there was no footage of a plane hitting the tower until several days later, Bush's recollection cannot be entirely accurate. However, it is important to recognize that his described thoughts and emotional responses would have been appropriate for anyone in his position, had they seen such footage. The simplest explanation of Bush's misremembering is that he saw plenty of video coverage of the attacks over the course of the days and weeks after 9/11. Since visual experiences are particularly noteworthy, it is plausible that the false memories were the result of misidentifying the relationship between the imagery he called to mind and the event during which he learned of the attacks.[15] He called to mind imagery he

[15] Some think that this and most other kinds of misremembering involve "source monitoring" errors i.e. failures to track the source of (part of) a representation. See Mitchell and Johnson (2000) and Neisser (1981) for discussion.

did see, but misidentified the source of those images. This led him misremember his own experience. His mind constructed a representation of something that one would expect to have happened, not what actually did happen.

Such cases are far from anomalous. Parallel patterns of misremembering have been observed for other significant events as well,[16] and in many cases, the results are striking. For example, 84% of U.K. respondents in one study "remembered" seeing footage of the 2005 bus bombing at Tavistock Square despite the fact that there was no such footage (Ost *et al*. (2008)). Similarly, 73% of U.S. respondents shared Bush's error in "remembering" having seen the first plane strike Tower 1 *on* September 11 (Pezdak (2003)).

The data is interesting because it illustrates the fact that normal human subjects *do* misremember personally significant events in their own lives, that such misremembering may be widespread, and that we cannot necessarily determine whether we are misremembering such things by reflection alone. The question is whether any of this is in tension with an epistemic theory of remembering.

For there to be any challenge, an opponent of G.E.T. would need these cases to motivate the conclusion that memory is somehow unreliable. However, it is not clear why this is the conclusion that should be drawn. Flashbulb memories are distinctive in a number of ways. They pertain to one's personal experiences and actions during a time in which mental resources are largely dedicated elsewhere (e.g. one primarily focuses on the significant event itself, its import in one's own life,

[16] For example, the death of Princess Diana (Ost *et al.* (2002)); The O.J. Simpson verdict (Schmalk *et al.* (2000)); the Challenger Disaster (Neisser and Harsch 1992: 9), the Reagan assassination attempt (Pilleman (1984)); and the assassinations of Martin Luther King Jr. and John F. Kennedy (Brown and Kulik, (1977))

etc. and not on one's *experience*). Given this, it is unclear what such cases reveal about the overall reliability of one's memory faculties.

Moreover, the data does not suggest that memory has led roughly 75% of people to be seriously confused about the gist of what transpired on, e.g. 9/11. Nor does it reveal anything in particular about the frequency with which any specific individual will make these kinds of memory errors. At most, it shows that, within a population of normal subjects, many (even most) may be wrong about some aspect(s) of their prior experiences. Though perhaps still striking because of the lack of correlation between subjects' confidence in the accuracy of their recollection, and the actual accuracy of their recollection, once thought of in this way, these results are not at all surprising and not at all a challenge to the idea that memory is reliable.

A more interesting moral is that such results provide a nice illustration of the fact that whether or not we remember the details of experiences in our lives depends, in part, on the information we have encountered between the experience and a subsequent recollection of it. This is precisely what one should expect if memory were to be a genuine source of knowledge as it is in the G.E.T. framework. It is simply an instance of more general phenomena that make any fallible source of knowledge susceptible to plausible but misleading testimony. That large percentages of subjects have been led astray by intervening "testimony" (whether acquired inter-personally or intra-personally) reveals only that in some special circumstances (at least those in which we form flashbulb memories) we are more likely than not to receive testimony that is misleading between the time of our

experience and our subsequent attempts at recollection. This does nothing to suggest that our memory faculties are unreliable.

It is also not clear that such cases would challenge G.E.T. if we granted, for the sake of argument, that they are evidence of some sort of lack of reliability. My opponent needs the cases to support the conclusion that memory is unreliable in a way that some instances of remembering will fail to be instances of knowledge. In these cases, we are struck by the fact that before such studies were done, we were unaware of this particular limitation of our memories. Now that we are aware, we plausibly question both whether we *know* the details of our flashbulb memory experiences, and whether we really *remember* such details at all. So, remembering and knowing seem to be equally challenged by this data (to the extent that either are challenged at all). Consequently, this kind of misremembering does not support the claim that memory is unreliable in a way that undermines G.E.T.

Consistency biases in memory

A second kind of case one might take to be evidence of memory's unreliability are those in which normal subjects have a tendency to misremember their prior attitudes, emotions, and experiences when they are inconsistent with their present attitudes, emotions, and experiences.[17] Such cases in which memory representations are biased towards being consistent with present mental states can be illustrated as follows:

Pain

[17] See e.g. Shanton, 2011 for one version of this argument, though she puts it to a different use.

Unfortunately, many medical procedures are unpleasant, even painful. In many cases, particularly painful experiences seem to stay with us in the sense that we naturally say things like, "I remember how painful that was. I hope to never go through it again." And it would be natural to think that our recall of the painfulness of prior medical procedures would track the intensity and duration of the pain experienced during those procedures. This thought appears to be mistaken.

Eich *et al*., 1985 examined regular participants in a pain management program. At an initial meeting, patients with histories of chronic headaches were asked to keep detailed pain diaries over the course of roughly one-week intervals. After the diaries were collected, participants also completed measures of their current pain intensity, and their recalled maximum, minimum, and usual levels of pain intensity during prior week.

The researchers found that patient memories for the physical pain experienced during a prior medical procedure were a function of their *present* level of pain. 76% of those experiencing significant pain at the time of recall overestimated their prior pain experience and 75% of those not experiencing significant pain at the time of recall underestimated their prior pain. In short, what patients seem to remember about their prior pain experiences is determined not just by those prior experiences, but also by their present experiences. In particular, they tend to skew their "memory" of past pain experience to be consistent with their current experience.[18]

[18] Redelmeier, Katz, and Kahneman (2003) and Redelmeier and Kahneman (1996) found related effects for patients undergoing painful medical procedures. Recollected prior pain experiences were largely a function of two things: (i) the intensity of the experienced pain during the final three

Emotions

Sometimes we experience upsetting events; sometimes they are so upsetting that our responses seem to stay with us in the sense that we say things like, "I remember how sad I was when such and such happened." And it would be natural to think that our recollection of those feelings depended on their intensity, duration, etc. This too appears to be mistaken.

Levine (1997:174) studied supporter reactions to Ross Perot's abrupt withdrawal from (and subsequent re-entry into) the race for president in 1992. She found that his supporters recall of their past emotional reactions were systematically distorted in parallel with their changing appraisals. Those that remained loyal to Perot throughout the election underestimated how sad they had felt when he first withdrew and overestimated their initial feelings of hopefulness. Those who changed their allegiance after his withdrawal but also later wished that Perot had been elected underestimated how angry they had been when he first withdrew from the race. And those who sustained their withdrawn support underestimated their prior feelings of sadness and hope. The upshot is that, while past emotions that were consistent with present appraisals tended to be remembered accurately or slightly overestimated, those that were inconsistent with present appraisals were significantly distorted.

Attitudes

Finally, similar trends have been observed for recollection of prior beliefs as well. In a widely cited study, Goethals and Reckman (1973) asked a sample of high

---

minutes of the procedure and (ii) the peak intensity of the pain. "Recalled" pain did not reflect the duration or aggregate intensity of the pain experienced during the procedure.

school students their opinions on a variety of social issues. They then invited participants back a couple of weeks later for further discussion of one particular issue (namely, whether "bussing" should be used to achieve racial balance in the nation's schools). They split the students into discussion groups in accordance with their previous responses (those in favor in one group and those opposed in another), and ensured that a confederate (student they had prepared with compelling counterarguments) initiated group discussions.

Subsequent measures showed that (i) the confederates' arguments did change people's beliefs; (ii) people did not take the arguments to have affected their beliefs (rather, they took the discussion to have reinforced their prior beliefs); (iii) people misremembered their prior beliefs as being consistent with their present beliefs.[19] The upshot is that, one's memory of one's prior attitudes doesn't always only depend on those attitudes, but is sometimes a function of one's present attitudes.

These three kinds of cases impressively demonstrate that one's present mental states play a much larger role in determining what we remember (or think we remember) about our past experiences, emotions, and beliefs than we might have thought. Again, the key question is whether any of this is inconsistent with G.E.T. An opponent might think they are because these cases seem to show that memory is unreliable. After all, one's (other) present mental states are surely irrelevant to the aim of accurately representing one's prior mental states (i.e. experiences, emotions, and beliefs).

[19] A control group demonstrated that subjects had not likely forgotten their prior beliefs before participating in the discussion.

There are at least two available responses. First, it is not clear that the claim that one's present mental states are irrelevant to the aim of accurately recollecting one's prior mental states is ultimately tenable. A plausible explanation of the existence of consistency biases is that one's present mental states are, on balance, a reliable indicator of one's prior mental states. They are by no means a perfect one, and some of the ways in which we can be misled by our present mental states are striking. However, this does not mean that they are unreliable indicators of past mental states overall. Consider the scope and variety of one's current beliefs. For example, I believe that my grocery store is on 38th Street; I believe that I like coffee; I believe that dislocating one's shoulder is painful, etc. These beliefs, and countless others, certainly do provide a good guide of what my past beliefs were.

In fact, present mental states often provide crucial testimony – sometimes they are the *only* available information – as to one's prior mental states. The fact that we find particular instances in which one's present mental states skew one's recall of one's past mental states, simply doesn't show that present mental states are irrelevant to the aim of recollecting prior mental states, and so doesn't show that memory is unreliable. To hold otherwise would be to endorse radical skepticism about our knowledge of our past.

Second, none of these cases suggest that memory is unreliable in a way that threatens to make remembering and knowing come apart. In all three cases, there is no temptation to judge that one remembers that *p* but does not know that *p*. Rather, they are all cases in which we judge that the subject fails to remember that *p*. We may find it surprising that we don't remember, exactly how painful something was,

exactly how we felt about certain events (perhaps these two aren't that surprising), and sometimes what we previously believed about certain topics. Our surprise does not indicate that memory is epistemically unreliable; it shows only that we don't remember as frequently as we might have thought we did.

In sum, I have argued that the kinds of misremembering observed in flashbulb memory cases and cases involving consistency biases do not demonstrate that memory is epistemically unreliable—that is, unreliable in a way that gets remembering and knowing to come apart (without yielding skepticism). Part of the reason for this is that in all of the cases at issue, the subjects clearly do not remember that *p*. An opponent would formulate a more interesting challenge using cases in which subjects clearly do remember that *p*. And then arguing that there is something that prevents this state of remembering from being formed in a manner conducive to knowledge. I turn to such cases next.

5.3.2 Vulnerable Memories

An opponent might instead argue that even when our memories do not fail us, G.E.T. is mistaken because we could *have very easily been wrong* about what we remember. One could motivate this by appealing to so called, 'implanted rich false memories' i.e. vividly detailed, emotionally laden, memories of fictitious events that can be easily generated in experimental and ordinary contexts. [20]

Researchers have consistently demonstrated that one can generate rich false memories in normal subjects simply by presenting them with credible, but

[20] See e.g. Shanton, 2011 for such an argument.

misleading, information about actual or putative to-be-remembered items.[21] For example, Frenda *et al.* (2013) found that half of their participants misremembered fictional political events when presented with doctored or misleadingly captioned photos of politicians. I will briefly elaborate:

Each participant was presented with photographic depictions of three true events and one of five target false events. After each viewing, participants' memories were probed using two measures. First, they were asked to select among the following options: (i) I remember seeing this; (ii) I don't remember seeing this but I remember it happening; (iii) I don't remember it; (iv) I have a different memory of how it happened. Second, they were given the option of elaborating by answering two open-ended questions: (v) How did you feel about this event at the time? (vi) Looking back, how do you feel about it today?

To illustrate, I include one of the false event presentations and some of the corresponding results below:

> Hillary Clinton's attack ad*:* Participants were shown a still image from a Republican advertisement featuring Barack Obama and Rev. Jeremiah Wright that was doctored to look like Hillary Clinton had approved it. The image as captioned as follows: "April 14, 2008: Trailing in the delegate count for the Democratic presidential nomination, Sen. Hillary Clinton airs an ad in Pennsylvania linking Sen. Barack Obama to the Rev. Jeremiah Wright. Under criticism, she pulls down the ad but wins the primary." In fact, Clinton never aired such an advertisement (Frenda *et al.* 2013: 281).

Some of the relevant results follow: (i) a large majority of participants (82%) reported remembering all three true events*;* (ii) 68% of participants "remembered" *seeing* the fictitious event on television with (iii) some elaborating on their false

[21] Misinformation research has focused on a variety of things we can remember (e.g. personal events in one's past, (specific and generic) public events like those discussed in the text, properties of oneself (e.g. being left-handed)) and dates back to at least to Loftus and Palmer (1974) and Goethals and Reckman (1973). The misinformation in these paradigms has also come from a variety of sources (e.g. participants' families, respected radio authorities, computer programs, doctored photos, guided imagination). See Wade *et al.*, 2007 and Loftus and Cahill, 2007: 413-25 for some review.

memories by reporting things like, "I thought it was a desperate move and it solidified my disgust with Mrs. Clinton as a candidate" (Frenda *et al*. 2013: 283).

The crucial point is not that we are probably making these kinds of errors all the time but rather that it is *tremendously easy to manipulate memory* in this way. Encountering plausible but misleading evidence of an event can give rise to the formation of false memories of that event including, memories of prior beliefs that one never held, related emotions that one never experienced, and so forth.

In response to prior cases, I have argued that the evidence only showed that, like any other imperfect source of knowledge, memory is subject to being led astray by misdleading information. This objection argues that this very feature of memory creates a problem for G.E.T. The central idea is that, even when memory manages to hit on the truth, it does so in a precarious manner. To put it another way, memory beliefs are not epistemically 'safe.'

Let us grant, for these purposes, that some version of safety is a necessary condition on knowledge.[22] I will argue that the cases used to illustrate the misinformation effect are consistent with such theses – In cases in which one remembers that *p*, one's true memory beliefs are safe. Consequently, the senses in which our true memories are vulnerable do not challenge G.E.T.

The misinformation effect paradigm shows that, in at least some circumstances, it is easy to convince someone that they remember something that is not the case. In the example case described above, researchers were able to convince many participants (68% of them) that they remembered seeing an attack

[22] Though, for some arguments against various versions of safety thesis see: Baumann, 2008; Bogardus, 2012; Comesaña, 2005; Neta and Rohrbaugh, 2004; Shanton, 2011; and Yamada, 2011.

ad that they did not in fact see. G.E.T.'s opponent needs cases of this sort to support the conclusion that true memory beliefs are not safe. To see why it does not, consider the following representative formulation of safety thesis:

> Safety: If a believer knows that *p*, then in nearly all, if not all, nearby possible worlds in which the believer forms the belief that *p* in the same way she does in the actual world, that belief is true. (Pritchard 2005, 163)

Using it, one could construct an argument against G.E.T. as follows. (Premise 1) G.E.T. entails that when *S* remembers that *p*, her belief that *p* is safe. (Premise 2) *S's* belief that *p* is safe only if, in nearly all nearby possible worlds in which she forms the belief that *p* in the same way as she does in the actual world, her belief that *p* is true. (Premise 3) It is not the case that in nearly every nearby possible world in which *S* forms the belief that *p* as she does in the actual world, her belief that *p* is true, for it is likely that in a world in which *S* is subject to misinformation with respect to *p*, she will form a false belief. (Conclusion 1) So, *S's* belief that *p*, when she remembers that *p*, is not safe. (Conclusion 2) So G.E.T. is false.

This argument does not succeed. My opponent wants to use the misinformation effect to show that true memory beliefs come out false in some nearby possible worlds. However, the cases do not show this. Rather, they show that, when presented with misleading evidence, subjects can be (easily) made to form different (false) beliefs. Consequently, premise 3 is simply confused and does not suggest that one's memory beliefs fail to be safe when one remembers that *p*.

Consider the attack ad case. Suppose that before the misinformation, *S* at least dispositionally believed (truly) that she does not remember seeing a Clinton ad

that attacked Obama's link to Rev. Wright. *That* belief is not false in nearby possible worlds in which she is exposed to misinformation. In worlds in which she is exposed to misinformation *S* forms a *different belief*: namely, that she does remember seeing a Clinton attack ad.

While this new belief is false, it is irrelevant. We can see why by considering an analogy involving visual illusions. Suppose that *S* sees two parallel line segments that are of equal length, and comes to believe that they are of equal length. Now consider a counterfactual situation in which she is shown two parallel lines that are of equal length, but which are made to appear to be of different lengths as in the Müller-Lyer illusion and comes to believe that they the line segments are not of equal length. In this counterfactual scenario, *S* is presented with misinformation and, as a result, forms a different and false belief. This counterfactual possibility in which a subject forms a false belief on the basis of misinformation does nothing to impugn *S's* knowledge when she sees that the lines of are equal length in the actual situation. Likewise, that it is easy to implant false memory beliefs via misinformation tells us nothing about cases in which one remembers that *p*.

5.4 Remembering as evidence part 2: *a priori* arguments

So far, I have responded to *a posteriori* objections to G.E.T. Some have argued, *a priori*, that G.E.T. mistakes the relationship between remembering and knowing. Their strategy is to present counterexamples to the following entailment thesis: if *S* remembers that *p* at *t*, then *S* knows that *p* at *t*. The thought is that, if there are clear possible cases in which one remembers that *p* at *t* without knowing that *p* at *t*, then

it cannot be the case that remembering is a species of knowledge. Instead, it is plausible that remembering is evidence that memory makes available to subjects.

The thought that motivates pursuit of counterexamples to the entailment thesis is that there is some sense or senses in which it is easier to remember something than it is to know it.[23] And one can isolate the relevant sense(s) in which remembering that *p* might be thought to be easier than knowing that *p* by examining the necessary conditions for knowledge.

At the very minimum, propositional knowledge is taken to have at least four necessary conditions:

S knows that *p* only if

(i) *p.*

(ii) *S* believes that *p.*

(iii) *S* is justified in believing that *p.*

(iv) *S's* belief is not *epistemically* lucky.

Most participants in this literature share commitment to the claim that remembering is factive. And this is reasonable as, arguably, one cannot remember, e.g. that she went to Paris as a child or that Paris is the capital of France, unless she actually did go to Paris and unless Paris actually is the capital of France, respectively. So the supposed discrepancy between knowledge and remembering must be in conditions (ii) – (iv).

I will consider example cases involving each condition and offer three responses to each case. First, putative counterexamples are hostage to the strength

[23] See Naylor (1986: 295), Lehrer and Richard (1975: 124), and Bernecker (2010, personal correspondence), and Galen Strawson (personal correspondence).

of the intuitions they elicit. The cases that have been offered in opposition to the entailment thesis are unusual and, at best, borderline instances of either remembering or knowing that *p*. Consequently, the intuitions that guide judgments about the cases are weak/varied and judgments are not conclusive. This limits the cases' force as apparent counterexamples. Second, once we isolate the considerations that guide our judgments about the cases, we see that those that tempt us to conclude that the subject remembers that *p* also tempt us to conclude that the subject knows that *p*; and those that tempt us to conclude that the subject fails to know that *p* also tempt us to conclude that the subject fails to remember that *p*. This may provide a partial explanation of why our judgments are so unstable. Third, I argue that once we isolate each relevant feature (belief, justification, luck) by formulating "limit cases" to test our assessments, there is little temptation to conclude that there are genuine counterexamples to the entailment.

The upshot is that opponents of G.E.T. have not offered genuine counterexample cases in support of the claim that it is in some relevant sense easier to remember something than it is to know it. Moreover, given the complex nature of the topic (it involves *both* memory and knowledge evaluations), and the reasons for which these cases fail to be counterexamples, it is unlikely that compelling cases are forthcoming. Consequently, the success or failure of theories of remembering turns on other considerations.

5.4.1 Belief

Sven Bernecker has recently argued that remembering that *p* does not require believing that *p,* and *ipso facto*, does not require knowing that *p*, by adapting a case from Malcolm (1963):

> [A]t t_2 S suddenly finds himself with the thought that he has been kidnapped when he was a small boy (at t_1). The idea that he has been kidnapped just pops into his head; it seems to come 'out of the blue'. S can't make sense of this idea and takes it to be merely imaginary. After all the likelihood of being kidnapped is rather low. What is more, the idea in question is inferentially isolated from the large body of inferentially integrated beliefs to which S has access. Nothing of what S knows or believes about his past connects with the idea that he has been kidnapped. But now suppose that, unbeknownst to S, it is in fact the case that he has been kidnapped. The flashbulb thought is an instance of propositional memory. Perhaps because of the terror of the experience S can't allow himself to even consider the possibility that he had been the victim of kidnapping but instead takes himself to be making it up. (Bernecker, 2010: 88)

According to Bernecker, this is an instance in which the subject remembers that he was kidnapped as a child. However, it would be wrong to say that *S* believes (or knows) at t_2 that he was kidnapped as a child because he would have to be presented with significant additional evidence before he would even reluctantly accept that such an event had occurred. Thus, the subject neither occurrently nor dispositionally believes that he was kidnapped when the thought 'pops' into his head (2010:89). To put it another way, one might take the case to show that memory provides subjects with evidence for belief but does not guarantee belief for one can (rightly or wrongly) reject the evidence as spurious. In contrast, in the case of knowledge, belief is *constitutive* and so the knowledge condition is subject to counterexample.

I will offer three responses to this case. First, it is a borderline case of remembering that *p*, at best. It may be that *S* represents a true proposition *p* as a result of something his memory processes have done. However, in the absence of a worked out theory of remembering, it would be quite a jump to conclude that he

thereby remembers that *p.* We might think that it is simply analogous to a jump from the premise that *S's* truly represents that *p* to the conclusion that *S* knows that *p*. Consequently, intuitions about the case are not particularly robust and reasonable people disagree about what conclusions can be drawn on its basis.

Second, the considerations that motivate ambivalence toward the case actually support pairing knowledge and remembering judgments together. The central motivation for thinking that *S* does not believe that *p* is that *S* takes no relevant kidnapping to have occurred and would need additional information to accept *p*; that is to say, *S* disputes *p*. We do not generally take subjects to remember facts that they actively dispute, and to the extent that we do, we take them to believe the proposition they claim to dispute. To see this, consider the following examples:

First, suppose that a young student is asked to name every national capital she can remember during a World Geography class. Suppose also that she successfully recalls the capitals of many nations in North and South America, Europe, and Asia. However, she notably omits Paris. Then suppose that when the possibility is brought to her attention, she actively disputes the claim that Paris is the capital of France. In the absence of any further behavioral evidence to the contrary, once our attention is drawn to her resistance, the question as to whether she remembers that Paris is the capital of France seems to be answered – she does not.

A plausible explanation is that unless one accepts *p*, one is in no position to do any of the things we generally associate with remembering it. For example, in addition to the error described above, a subject that disputes that Paris is the capital

of France is in no position to guide her own behaviors with respect to the French capital (e.g. to purchase a ticket to the capital of France should she want to visit it), on the basis of anything she remembers. Similarly, we might think, in the case described, the subject that disputes the fact that he was kidnapped as a child is in no position assert that he was kidnapped on the basis of anything he remembers; nor is he in a position to guide any of his other behaviors with respect to that event. Consequently, he does not really remember that he was kidnapped at all. In this way, the absence of belief pushes us to conclude that one does not remember. [24]

On the other hand one can sometimes "refuse to believe that p", when one probably does, in fact, believe that p. For example, one may "refuse to believe" that one's partner is cheating, but the very energy with which the proposition is disclaimed may be evidence of an underlying, and unwelcome belief in it. And so, we might think that, like one's psychological trauma may lead one to deny believing that a partner is cheating, when, one in fact believes the partner is, so too may the 'terror of the experience' lead S to not 'allow himself to even consider the possibility that he had been the victim of kidnapping' when in fact, he does believe it after all. If so, a judgment that S remembers that *p* is consonant with the judgment that he really does (deep down) believe that *p*.

The important point is that whichever way we go, once we isolate the considerations that guide our judgments about whether a subject holds, or does not hold a particular *belief*, we see that the considerations that lead to accepting or

[24] The same points apply in a case in which *S* represents, but is agnostic about, *p* at t_2. The only evidence that it is an instance of remembering appears to be that *S's* representation emerges from memory processes. *S* would still be unable to do all of the things we normally associate with remembering (e.g. make decisions on the basis of remembering that *p*, etc.)

denying that S remembers that *p*, also lead to accepting or denying that he believes that *p*. Consequently, this putative case of remembering without believing is not compelling. However, even if one were to find a particularly convincing case of remembering that *p* without believing that *p*, the opponent of the entailment would simply face a new challenge. Namely, such a convincing case would highlight a need to revisit the role of belief in knowledge.

To oppose G.E.T. one would have to show that despite these particularly compelling kinds of cases, knowledge actually requires *belief* and not whatever virtuous attitudinal state is involved in the remembering state. To put it another way, the presentation of a particularly compelling instance of remembering that *p*, e.g. one in which the subject is positioned to make decisions regarding *p*, to assert that *p* on the basis of her remembering, and so forth, the opponent would have to convince us that knowledge would nevertheless require something further—'belief' (as characterized by a technical definition) which fails to be satisfied by the remembering state.[25]

Finally, if we reformulate the case so as to remove the attitudinal elements that push in either direction, i.e. if we formulate a 'limit case' in which we are neither tempted to conclude that the subject entertains a pro-attitude with respect to *p*, nor disputes the claim that *p*, we are left in no position to make any judgments about whether he remembers that *p*. We *can* conclude that he somehow has some kind of mental state with *p* as its content and that that mental state is somehow the

[25] This problematic result is compounded by the fact that some productive accounts of knowledge already question whether belief essential to knowledge (see e.g. Farkas, 2014; Meyers-Schutz and Schwitzgebel, 2013; Williamson, 2000; and Black, 1971).

causal product of some prior events (including the event that makes it the case that *p*) and some memory processing. Saying anything further would require considerable theoretical commitments. The upshot is that whether or not a subject believes that *p* is not plausibly a source of counterexample to the entailment between *S* remembers that *p* and *S* knows that *p* and so not plausibly a source of counterexample to G.E.T.

5.4.2 Justification

An opponent might also offer cases meant to show that a subject can remember that *p* despite not being justified in believing that *p*. There are two central ways to do this.

First, some philosophers think that evidence one does not possess can undermine one's justification for believing that *p*. A variant of a case from chapter 2 (p. 32) can be used to examine the issue in memory contexts (adapted from Harman (1976) *cf*. Dretske and Yourgrau (1983)):

> On Tuesday, everyone knows that the king is dead. They know this because they all watched the funeral on television, read his obituary in the papers, and so forth. On Wednesday, the government undertakes a massive deception—telling everyone that the king did not really die; the previous "news" was all an elaborate hoax. They even support this claim by orchestrating a news conference at which the king "appears" to assure the people that reports of this death were greatly exaggerated. Suppose further that everyone but Clyde gets the message. They all now believe that the king is alive and well. Clyde, however, has been cloistered in his study working on his thesis since Tuesday night. As a result, on Thursday, he still believes that the king is dead, and he believes it on the same grounds that everyone (originally) had for believing this.

As Dretske and Yourgrau observe (1983:357), this is a peculiar case and there is more than one reasonable way to respond to it. On one side, it is plausible that Clyde remembers that the king is dead. He certainly *seems* to remember it. After all, he formed his true belief on the basis of reliable information sources and has not

ostensibly forgotten anything. However, the very same considerations make it plausible that Clyde also knows that the King is dead. He formed the true belief in an epistemically suitable way, and he has no evidence (misleading or otherwise) that he is somehow mistaken. Consequently, it is difficult to see why we should find fault with the epistemic status of Clyde's belief simply because he was not in the unfortunate position of having received misleading and false information about the king. That everyone else fails to know that the king is dead when they receive the misleading evidence only suggests that, for a time, Clyde was the only lay-citizen to know that the king was dead. And so the considerations that point in favor of judging that the subject remembers that *p*, also point in favor of judging that he knows that *p*.

Alternatively, one might be moved by the fact that apparent defeaters are available to the members of Clyde's community and think that, in spite of having ostensibly forgotten nothing he knew on Tuesday, Clyde can no longer know that the king is dead because his knowledge would be based on the ostensibly defeated evidence. If one finds this position is compelling, she might also be tempted to think that the case constitutes a counterexample to the knowledge condition. To do so, she need only show that Clyde hasn't merely seemed to have forgotten nothing; he has genuinely forgotten nothing. That is, she must motivate the claim that Clyde's remembering that the king is dead persists in the face of defeating evidence that he fails to possess, but that nevertheless undermines his knowledge.

It is not clear that this position is particularly plausible beyond a first blush. The fact that Clyde has not ostensibly forgotten anything he knew on Tuesday is the

only support for the claim that he remembers that the king is dead. However, seeming to have not forgotten anything need not guarantee that one now remembers everything one once did for, as Bernecker (2010: 198-201) and Tye (1998) have observed, there is more than one way to forget things. We are most familiar with memory failures that are the result of some internal failure of the memory system. However, some memory failures are the result of environmental factors outside of the subject's internal cognitive systems.

The paradigmatic instances of such forgetting involve cases of slow-switching either across worlds or across time in our own world but the analogy with content externalism need not be perfect.[26] It simply highlights the fact that whether or not a subject remembers that *p* is partially determined by the external environment around her *at the time of recall*. While it may be that Clyde has not undergone any forgetting as a result of internal failures of his cognitive system, one would need an antecedent understanding of the relationship between remembering and putative defeating evidence to determine whether Clyde continues to remember that the king is dead on Thursday.

Moreover, consider how we would evaluate Clyde in a counterfactual circumstance in which the king really was alive. Uncontroversially, we would resist attributing knowledge that the king is dead. After all, Clyde's belief would be false, and we would be able to point to the relevant evidence against his position as justification. Of course, the very same considerations would give us reason to reject

[26] See Tye (1998: 87) for discussion.

the claim that Clyde remembers that the king is dead.[27] In contrast, when we consider the actual case in which the king is in fact dead, it seems just as reasonable for us to say that Clyde remembers that the king is dead, citing our own evidence for this, and to use the very same evidence to justify the claim that he does in fact continue to know that the king is dead. So, just as the considerations in favor of the judgment that Clyde remembers, also favored the judgment that Clyde knows, the considerations that favor the judgment that Clyde *does not know* also support the judgment that he does not remember, that the king is dead.

A second, and perhaps less controversial, kind of case is as follows. Most philosophers accept that a subject is not justified in believing that p, if *she is in possession* of defeating evidence. To see the import in memory contexts consider the following case adapted from Bernecker (2010: 78).[28]

> Suppose that at t_1, *S* learns that the Colosseum was completed in [the year] 80 (*p*). Suppose further that at t_2, she gets some plausible, but ultimately misleading, evidence that it was completed in [the year] 90. Suppose that although she continues to believe that *p*, she cannot rule this new evidence out, and so cannot know that *p*.

Bernecker concludes that this is a counterexample to the entailment between remembering and knowing that *p* because he maintains that despite the defeating evidence, *S* remembers that the Colosseum was completed in the year 80. However, it is difficult to see exactly why. Given the story, and in particular, the fact that *S* is aware of the (misleading) evidence against *p*, it's unclear why she continues to

[27] We could of course allow that Clyde *seemed* to remember that the king was dead; that he clearly *does* remember that he was *certain* the king was dead on Tuesday; that he saw what seemed to be a funeral service; and so forth, but it is also reasonable to think that he knows these things as well.
[28] See Dretske and Yourgrau (1983); Ginet (1988: 160); Naylor (1986:298); Saunders (1965: 282-3); and Harman (1973) for similar cases.

believe that *p* at all, as others have already pointed out.[29] If she were able discount this misleading evidence for some reason or other, then one could understand why she would believe that *p*. Of course, it would also then be plausible to say that she still knows that *p* as well, as it would no longer be clear that there are defeaters. On the other hand, if she genuinely cannot rule out conflicting evidence, it seems that her memory has betrayed her. Given the available information, it should have produced some other attitudinal state, e.g. that the Colosseum was completed in either 80 or 90. At most, we can conclude that memory has provided the subject with a true belief that *p*, and it would be a jump to move from this to the conclusion that *S remembers* that *p*.

To see why, consider again the following limit case, i.e. case in which the subject uncontroversially lacks justification for the belief that *p*, despite her memory producing the true belief that *p* (cf. p. 57). Suppose that one morning I simply wake up with the belief that there is life on Planet G. I received no evidence in favor of this claim, and I don't particularly care whether or not there is life elsewhere. I believe it for the sake of believing something. Now suppose that there is life on Planet G and later, in virtue of the usual memory processes, I find myself again with the occurrent belief that there is life on Planet G. In such a case, I would have a true belief formed on the basis of familiar memory processes, though, on any conception of epistemic justification, there is no point at which my belief is justified in any way, and so it is profoundly implausible that I should know that there is life on Planet G. Of course, it is also profoundly implausible that I remember that there is life on Planet G on this

[29] See e.g. Phillips (2012) and Adams (2011).

basis. Remembering is one way to be connected to the world, and remembering facts about the world depends on *some* relevant relation to those facts. This kind of case shows that genuine failure to be justified in believing that *p* plausibly undermines remembering and knowing that *p* alike.[30]

To summarize, judgments about this case may vary because the considerations that lead us to think that there are apparent defeaters either undermine the genuine success of the memory processes that produce the belief that *p* and so undermine the subject's ability genuinely remember that *p*, or they fail to undermine the success of those processes, and fail to undermine one's knowledge that *p*. Moreover, once we isolate the purported source of the counterexample (the absence of justification) by formulating a limit case, we have no reason to conclude that a true, but unjustified, memory belief should count as an instance of remembering. The upshot is that like belief, justification will not plausibly be the source of a genuine counterexample to G.E.T.

5.4.3 Epistemic Luck

The last source of potential counterexample is epistemic luck. Since at least Gettier's seminal 1963 paper, it has been widely accepted that one can have a justified true belief that fails to be knowledge because it is only a matter of luck that the belief is true. To examine luck's effect in memory contexts I will examine two cases. First, consider again the following case adapted from Russell (1948):

Russell's Clock

[30] See Audi (2010:64) for discussion of a few other cases in which a true belief that *p*, fails to be instances of remembering that *p*, despite being produced by one's memory processes.

> Suppose that one morning Jo comes in to class and looks up at the clock to check the time. It reads exactly 8:22, and as the clock had never led Jo astray before, she is justified in forming beliefs about the current time on the basis of its testimony. She does and comes to believe truly that it is 8:22. Suppose further that she gets the time right by mere luck for the clock had stopped working at exactly 8:22 the previous night.

Since it is a matter of dumb luck that Jo formed a true belief (there is no relevant connection between what the clock 'reads' and the actual time; consequently, had she arrived a moment later or earlier, Jo would have formed a false belief on the very same basis) most accept that Jo fails to know that it is 8:22 despite having a justified true belief. Now suppose that later that night, Jo tries to recall when she made it to class and on the basis of her prior belief and functioning memory processes forms the justified true belief that she came in at 8:22.

The only motivation for the claim that Jo remembers that she came in at 8:22 is that she has a true belief that depends on memory processes in most of the usual ways. Of course, the claim that one has a true memory belief that depends on memory processes in most of the usual ways, on its own, also supports the conclusion that one knows that *p*. More must be said to motivate both memory and knowledge judgments. In this case, the events that make belief true are entirely causally isolated from the subject's cognitive systems and highlighting this and the corresponding lucky nature of the true belief seems to me to undermine both remembering and knowing.

It should be no surprise that a memory belief that *p* that is only luckily true should fail to constitute remembering that *p* as the same holds in other domains as well. For example, one who forms a true perceptual belief that there is a sheep in a field does not see that there is a sheep in the field if the actual sheep is occluded by a

sheep façade.[31] Nor does one know *a priori* that *p* on the basis of making several fortuitous errors in reasoning that yield the true belief that *p*.

In sum, the considerations that point in favor of judging that a subject remembers that *p* in genuine Gettier cases (namely, true memory beliefs) are the same considerations that point in favor of judging that she knows that *p* in lucky circumstances. And the considerations that point against judging that she knows that *p* (causal isolation and corresponding epistemic luck) point against judging that she remembers that *p* as well.

Fake Barns

A second strategy for undermining the relationship between remembering and knowing can be derived from fake barn scenarios as follows.[32] Suppose that while driving through fake barn country, *S* just so happened to look up at the right moment to see the only real barn for miles. She could, on that basis, form the justified true belief that there is a barn off to the left. Of course, given that she is merely lucky to have her belief be true (a moment later and the very same type of belief would have been false for she would have been viewing a barn façade), many maintain that she does not know that there is a barn off to the left. Now suppose that when reminiscing about her trip at t_2, *S* recalls the barn off to her left on the basis of her prior experience and related belief. If she goes on to believe that there was a barn off to the left and, in so doing, *S* genuinely *remembers* that there was a barn off to the left, many would take this to show that one can remember that *p*

[31] The case is adapted from Chisholm, 1966. It is also worth mentioning that the same points apply to subjects that undergo a veridical hallucination (*cf*. Lewis, 1980) as of a sheep in the field.
[32] Ginet's "Fake Barn" style cases are courtesy of Goldman (1976: 772-73).

without knowing that *p*. After all, it is rather difficult to see how a belief that failed to constitute knowledge at t_1 could come to constitute knowledge at t_2 simply by virtue of having been stored and tokened again later.[33]

The proposal offered here has two avenues for response. First, one might deny that the subject remembers that there was a barn off to her left. If she does not remember that there was a barn, then it does not matter that she does not know that there was a barn. Second, one might deny that the kind of luck involved in this context undermines knowledge. I consider each in turn.

First, given the strength of the argument in favor of the view that remembering that *p* is a species of knowing that *p*, one might try to simply characterize the fake barn case so as to make it even remotely plausible that, given that the subject does not know that there was a barn, we have reason to believe that she doesn't really remember it either.[34] One possible way of doing so is as follows. When one successfully sees a barn, she becomes aware of, and has knowledge of, it. In normal circumstances, the visual information can be unproblematically transferred to other psychological systems, and in particular, those associated with cognition. As a result, she can come to see (and know) *that* there is a barn before her. In fake barn country, the unusual circumstances create a problem for the transfer of visual information into cognition. The subject sees the barn, but it is less clear whether she can be understood to see *that* there is a barn before her. While it is true that she has the conceptual resources necessary to see that there is a barn before her 'seeing-that', one might think, is subject to normative considerations like

[33]McGrath (2007: 19-22) for more discussion of this "epistemic boost problem."
[34] See Ranalli (2014) and French (2013) for discussions of this strategy for perceptual cases.

reliability. And in fake barn country, the normally reliable transition from lower level perceptual information to cognitive information is not always reliable. This, one might think, suggests that it is possible that the subject only *seems* to see that there is a barn before her, for she is merely lucky to have had her visual processing produce a true belief. In effect, though the visual system may have no internal flaws, the context in fake barn cases compromises the integrity of perceptual processing.

Similarly at t_2, the subject's memory processes are doing quite well. When she recalls (presumably imagistically) the barn off to her left, she remembers it, she is aware of it (in the past) and she thereby has knowledge of it. In normal circumstances, the mnemonic information can be unproblematically transferred to other psychological systems, and in particular, those associated with cognition. As a result, she can come to remember *that* there was a barn before her. Having been acquired in fake barn country, the unusual circumstances create a problem for the transfer of mnemonic information into cognition, just as they did initially in the perceptual case. We might conclude then, that although the subject remembers the barn, it is not clear that she remembers *that* there was a barn before her, even if she happens to form the true belief as a result of memory processes.[35] Just as with perception, contextual factors have plausibly compromised cognitive processing in terms of the transition from objectual mental state (remembering the barn) to corresponding propositional mental state (remembering that there was a barn).

My own preference is to take the second option - to deny that the kind of luck involved in this case undermines knowledge. One might do so by drawing the

[35] See Audi (2011: chapter 3) for some additional discussion of cases in which a true belief is formed as the result of memory processes but nevertheless fails to be something one has remembered.

following distinction: (i) *S* gets it right that *p* by luck because of *S's* circumstances; (ii) *S* gets it right that *p* by luck because of the basis from which *S* believes that *p*.[36]

This purported counterexample relies on luck of the first form, and it is not clear that this kind of luck undermines knowledge. Given that *S* was in fake barn country, it is a matter of luck that she saw a barn. Of course, it is often a matter of luck that we see what we see. For example, it is a matter of luck that I opened the newspaper to the sports page and saw the score of last night's game. Nevertheless, in doing so, I certainly came to know the score of last night's game. Similarly, though it is a matter of luck that *S* saw a barn, given that she saw a barn, she came to know that there was a barn. Once this is established there is no problem in concluding that *S* both remembers that there was a barn and knows that there was a barn. It is not at all a matter of luck that *S* is right in her belief, *given the basis on which she forms the belief* - her perception (and subsequent remembering) that there is a barn.

The upshot of this discussion is that G.E.T. is not easily subject to counterexample cases on the basis of discrepancies between remembering that *p* and knowing that *p* in terms of belief, justification, or epistemic luck. The cases are by their nature peculiar and borderline instances of remembering/knowing. Consequently, intuitive judgments about the cases are varied. Moreover, once we see that, in each case, the considerations in favor of judging that one remembers that *p*, also favor judging that one knows that *p*, and vice versa, there is little reason to think that compelling counterexample cases are forthcoming.

[36] See Ranalli (2014) for a related discussion in the context in perception. See Pritchard (2005) and Engel (1992), for versions of this distinction.

5.5 Conclusion

In sections 5.1 and 5.2, we saw that G.E.T. is consonant with the characteristics and purposes of human memory processing. In section 5.3, I argued that empirical research on the reliability and vulnerability of memories does nothing to diminish the epistemic status of remembering. Finally, the discussion in section 5.4, showed that remembering that *p* and knowing that *p* bear the same demands with respect to belief, justification, and luck. The upshot is that memory is not epistemically derivative as many have thought. It is a primary epistemic source and *remembering* is a species of knowledge.

Bibliography

Adams, Fred 2011: 'Husker du?' *Philosophical Studies*, 153, pp. 81-94.

Alea and Bluck (2003) "Why are you telling me that? A conceptual model of the social function of autobiographical memory" *Memory 11*, 165-178.

American Psychological Association (1994) Diagnostics and Statistics Manual (IV)

Anderson, M. (2007) "The massive redeployment hypothesis and the functional topography of the brain", *Philosophical Psychology,* 21(2), 143-174.

- (2010) "Neural reuse: a fundamental organizational principle of the brain," *Behavioral and Brain Sciences,* 33(4), 245-313.

Annis, D. (1980) "Memory and Justification", *Philosophy and Phenomenological Research* 40: 324-33.

Anscombe, G. (1981) "Memory, 'Experience' and Causation", in her *Collected Philosophical Papers*, ii. *Metaphysics and Philosophy of Mind*. Oxford: Blackwell, 103-119.

Aristotle (circa 350bc) *De Memoria et Reminiscentia*

Armstrong, J. & Stanley, J. (2010) "Singular Thoughts and Singular Propositions" *Philosophical Studies* Wednesday, March 10, 2010 (DOI 10.1007/s11098-010-9532-1)

Audi, R. (2010) Epistemology: A Contemporary Introduction to the Theory of Knowledge, Third edition, London: Routledge.

- (1995) "Memorial Justification" *Philosophical Topics* 23:31-45
- (1997), "The Place of Testimony in the Fabric of Knowledge and Justification," *American Philosophical Quarterly* 34: 405-22

– (2001) “An Internalist Theory of Normative Grounds” *Philosophical Topics 29* 19-46
– (2003) *Epistemology: A Contemporary Introduction to the Theory of Knowledge*. London: Routledge
– (2011) *Epistemology: a Contemporary Introduction to the Theory of Knowledge 3rd edition*. Routledge, New York.

Augustine (1991:191 (circa a.d. 397)) *Confessions*

Ayer, A. (1956) *The Problem of Knowledge*. Harmondsworth: Penguin.

Bach, K. (1987) *Thought and Reference* Oxford: Clarendon Press.
– (2010) “Getting a Thing into a Thought” in *New Essays on Singular Thought,* Robin Jeshion, ed. Oxford University Press

Barsalou, L. (1988) “The Content and Organization of Autobiographical Memories” in U. Neisser and E. Winograd (eds.) *Remembering Reconsidered: Ecological and Traditional Approaches to the Study of Memory*. Cambridge University Press pp.193-244

Bartlett. F.C. (1932) *Remembering: A Study in Experimental and Social Psychology,* Cambridge University Press

Baumann, P. (2008), “Is Knowledge Safe?”, *American Philosophical Quarterly,* 45 (1): 19-30.

Bechtel, W. (2008) *Mental Mechanisms: Philosophical perspectives on cognitive neuroscience* London: Routledge

Becker, J.T. and Morris, R. G. (1999), “Working Memory(s)”, *Brain and Cognition,* 41(1): 1-8.

Benjamin, B. (1956), "Remembering", *Mind*, 65 (July): 312-331.

Bentall, R. (1990) "The Illusion of Reality: A review and Integration of Psychological Research on Hallucinations" *Psychological Bulletin* 107: 82-95

Bergson, H. (1896) *Matter and Memory* trans. N.M. Paul and W.S. Palmer (1994), New York: Zone Books

Bernecker, S. (2007), "Remembering without Knowing," *Australasian Journal of Philosophy* 85, pp. 137–156.

- (2008) *The Metaphysics of Memory* Springer
- (2010) *Memory: a philosophical study* Oxford University Press
- (2011) "Memory Knowledge" in S. Bernecker and D. Pritchard (eds.) *The Routledge Companion to Epistemology* Taylor and Francis. Pp. 326-334

Black, C. (1971) "Knowledge without Belief", *Analysis*, 31(5): 152-8.

Blume, W.T., Girvin, J.P. and Sternerson, P. (1993) "Temporal Neocortical Role in Ictal Experiential Phenomena" Ann Neural (33: 105-7)

Boer, S. and Lycan, W. (1986) *Knowing Who, Cambridge:* MIT Press.

Bogardus, T. (2012) "Knowledge Under Threat", *Philosophy and Phenomenological Research,* 86 (1): 289-303.

Bower, G. (2000) "A Brief History of Memory Research", in E. Tulving and F.I.M. Craik (eds.) *The Oxford Handbook of Memory*. Oxford: Oxford University Press, 3-32.

Boyer, P. (2009) "Extending the range of adaptive misbelief: Memory 'distortions' as functional features." *Behavioral and Brain Sciences*, 32(6), 513-4.

Brain, W. R. (1962) *Diseases of the Nervous System*, London: Oxford University Press

Brewer, B. (1996) “Internalism and Perceptual Knowledge” *European Journal of Philosophy,* 4: 259-75

– (1999) *Perception and Reason* Oxford: Clarendon Press.

Brewer, W. (1986) “What is Autobiographical Memory?” in D.C. Rubin (ed.) *Autobiographical Memory* Cambridge University Press pp 25-49

Broad C.D. (1925) *The Mind and Its Place in Nature*. London: Kegan Paul, Trench, Trubner.

Broadbent, D. (1971) *Decision and Stress*, New York: Academic Press

Brown, R., and Kulik. J. "Flashbulb memories." *Cognition* 5.1 (1977): 73-99.

Buckner, R. (2007) “Integrative comments Memory systems: An incentive, not an endpoint” in *Science of Memory: Concepts*, Eds. H. Roediger, Y. Dudai, and S. Fitzpatrick, Oxford University Press, pp. 359-364

Burge, T. (1993) “Content Preservation” *Philosophical Issues* 6 (4): 271-300

– (1997) “Interlocution, Perception, and Memory” *Philosophical Studies* 86: 21-47

– (2003) “Memory and Persons” *The Philosophical Review* 112 (3): 289-337

Byrne, A. (2010) “Recollection, perception, imagination” *Philosophical Studies* Springer

Campbell, J. (1994) *Past, Space, and Self* Cambridge MA: MIT Press

– (1997) "The Structure of Time in Autobiographical Memory", *European Journal of Philosophy* 5 (1997), 105-118.

– (2002) *Reference and Consciousness* Oxford University Press

Campbell, S. (2006) “Our faithfulness to the past: Reconstructing memory value”. *Philosophical Psychology* 19 (3): 361 – 380.

Carroll, B. and Carroll, T. (2005) “Catatonia is the Rosetta Stone of Psychosis” Behavioral and Brain Sciences 28:6 759-760

Casey, E. (1976) “Comparative Phenomenology of Mental Activity: Memory, Hallucination, and Fantasy Contrasted with Imagination” *Research in Phenomenology,* 6, 1-25

Cassam, Q. (2007) “Ways of knowing”, *Proceedings of the Aristotelian Society*, *107*, 339–358.

Chisholm, R. (1966), *Theory of Knowledge,* Englewood Cliffs, N.J., Prentice-Hall.

Cohen, S. (1999), “Contextualism, skepticism, and the structure of reasons”. *Philosophical Perspectives* 13 (s13):57-89.

Cohen, N. & Squire, L. (1980) “Preserved learning and retention of pattern analyzing skill in amnesia: dissociation of knowing how and knowing that”, *Science,* 210, 207-209.

Collerton, D., Perry, E. and McKeith, I. (2005) “Why People See Things That are not There: A Novel Perception and Attention Deficit Model for Recurrent Complex Visual Hallucinations” *Behavioral and Brain Sciences* 28 737-757

Comesaña, J. (2005) “Unsafe Knowledge”, *Synthese,* 146(3): 395-404.

Conway, M. (2007) “Remembering: a process and a state” in *Science of Memory: Concepts*, Eds. H. Roediger, Y. Dudai, and S. Fitzpatrick, Oxford University Press pp237-242.

Craig, E. (1990), *Knowledge and the State of Nature: An Essay in Conceptual Synthesis,* Oxford University Press.

Crane, T. (2009) "Is perception a propositional attitude?" *The Philosophical Quarterly*, 59: 452-469.

- (2011a),"The Problem of Perception", *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/perception-problem/>.
- (2011b) "The Singularity of Singular Thought" *Aristotelian Society Supplementary Volume* 85 (1): 21-43

Currie, G. (2000) "Imagination, Delusion and Hallucination" in *Pathologies of Belief (Readings in Mind and Language),* eds. M. Coltheart and M. Davies Blackwell.

Dancy, J. (1985). *An Introduction to Contemporary Epistemology*. Oxford: Blackwell.

Davies, M. and Coltheart, M. (2000) "Introduction: Pathologies of Belief" in *Pathologies of Belief (Readings in Mind and Language),* eds. M. Coltheart and M. Davies, Blackwell.

De Brigard, F. (2012) "Predictive memory and the surprising gap." *Frontiers in Psychology*, (3), 420.

- (2014) "Is memory for remembering? Recollection as a form of episodic hypothetical thinking," *Synthese, 191:*155-185.

Debus, D. (2008) "Experiencing the Past: A Relational Account of Recollective Memory" *dialectica* 62 405-432

- (2010) 'Accounting for Epistemic Relevance: A New Problem for the Causal Theory of Memory', in: *American Philosophical Quarterly* **47**:1 (2010), 17-29

Dennett, D. & Westbury, C. (2000) "Mining the Past to Construct the Future: Memory and belief as forms of knowledge" in D. Schacter and E. Scarry (eds.) *Memory, Brain, and Belief,* Harvard University Press: 11-32.

Dennett, D. (1991) *Consciousness Explained* Back Bay Books

DeRose, K. (1992), "Contextualism and Knowledge Attributions", *Philosophy and Phenomenological Research,* 52 (4): 913-929.

– (2002) "Assertion, Knowledge, and Context", *Philosophical Review,* 111 (2): 167-203.

Deutscher, M. (1989) "Remembering 'Remembering'" in J. Heil (ed) Cause, Mind, and Reality, Dordrecht: Reidel, 53-72

Dokic, Jerome (2001), 'Is Memory Purely Preservative?', in C. Hoerl and T. McCormack (eds) *Time and Memory*, Oxford: Oxford University Press, 2001, pp. 213–232.

Donnellan, K. (1979) "The Contingent A Priori and Rigid Designators" in French, Uehling, and Wettstein, eds., *Contemporary Perspectives in the Philosophy of Language*, Minneapolis: University of Minnesota Press, 45-60.

Dopkins, S., Sargent, J., Ngo, CT (2013) "Exploring the Mental Representation Underlying Familiarity Assessment", *Canadian Journal of Experimental Psychology*, June: 67(2) 140-52

Dorsch, F. (2010) "The Unity of Hallucinations" Phenomenology and Cognitive Science 9:171-191

Dretske, F. (1981) Knowledge and the Flow of Information, Cambridge, MA MIT

Dretske, F. and Yourgrau, P. "Lost Knowledge" *The Journal of Philosophy* Vol. 80 No. 6: 356-67

Dudai, Y. (2002), *Memory from A to Z. Keywords, Concepts, and Beyond.* Oxford: Oxford University Press.

Dummett, M. 'Testimony and Memory', in his Seas of Language, Oxford: OUP, 1993, p. 421

Eich, E. Reeves, J. & Katz, R. (1985) "Anesthesia, amnesia, and the memory/ awareness distinction", *Anesth. Analg.* 64: 1143-48.

Engel, M. (1992) "Is Epistemic Luck Compatible with Knowledge?" *Southern Journal of Philosophy,* 30, 59-75.

Engel, S. (1999) *Context Is Everything: The Nature of Memory*, New York: W.H. Freeman and Company.

Evans, G. (1982) *The Varieties of Reference* Oxford: Clarendon Press.

Farkas, K. (2014) "Belief May Not be a Necessary Condition for Knowledge", *Erkenntnis,* DOI 10.1007/s10670-014-9620-2

Fish, W. (2009) *Perception, Hallucination, and Illusion* Oxford University Press

French, C. (2013), "Perceptual experience and seeing that P", *Synthese*, *190*(10), 1735–1751.

Frenda, S., Knowles, E., Saletan, W. & Loftus, E. (2013) "False memories of fabricated political events", *Journal of Experimental Social Psychology,* 49: 280-286.

Freud, S. (1925). "A Note Upon the 'Mystic Writing-Pad'". *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Volume XIX (1923-1925): The Ego and the Id and Other Works, 225-232

– (1959) *The Standard Edition of the Complete Psychological Works of Sigmund Freud* Volume IX

Furlong E. (1951) *A Study in Memory: A Philosophical Essay*. London: Thomas Nelson

Gall, F. (1835) "On the Functions of the Brain and Each of Its Parts: With Observations on the Possibility of Determining the Instincts, Propensities, and Talents, or the Moral and Intellectual Dispositions of Men and Animals, by the Configuration of the Brain and Head, Volume 1" Marsh, Capen & Lyon.

Garcia-Carpintero, M. (2004), "Assertion and the Semantics of Force-Markers", in C. Bianchi (ed.), *The Semantics Pragmatics Distinction.* Csli. 133-166.

Ginet, C. (1988) "Memory Knowledge", in G.H.R Parkinson (ed.), *The Handbook of Western Philosophy.* New York: Macmillan, 159-78.

Goethals, G. and Reckman, R. (1973) "The perception of consistency in attitudes", *Journal of Experimental Social Psychology,* 9(6): 491-501.

Goldman, A. (1967) "A Causal Theory of Knowing", *Journal of Philosophy,* 64 (12): 357-372.

– (1976). "Discrimination and Perceptual Knowledge." *The Journal of Philosophy*, 73: 771–791.

Greenberg, D. (2004) "President Bush's False 'Flashbulb' Memory of 9/11/01", *Applied Cognitive Psychology,* 18: 363-370.

Grice, H. (1941) "Personal Identity", *Mind 50:* 330-50.

– (1961) "The Causal Theory of Perception, Part 1", *Proceedingsof the Aristotelian Society,* 121:121-152.

Harman, G. (1973) *Thought*, Princeton, New Jersey; Princeton University Press

– (1986), *Change in View,* Cambridge, MIT Press.
– (1990) "The Intrinsic Quality of Experience," in *Philosophical Perspectives*, 4, J. Tomberlin, ed., Northridge: Ridgeview Publishing Company.

Hasselmo, M. (2012) *How We Remember: Brain Mechanisms of Episodic Memory*, Cambridge, MA, MIT Press.

Hawthorne, J. (2004) *Knowledge and Lotteries,* Oxford University Press.

Hawthorne, J. and Manley, D. (2012) *The Reference Book,* Oxford University Press

Hazlett, A. (2010) 'The Myth of Factive Verbs', *Philosophy and Phenomenological Research* 80: 3 pp 497–522

Henderson, D. (2009) "Motivated Contextualism", *Philosophical Studies,* 142 (1): 119-131.
– (2011) "Gate-Keeping Contextualism", *Episteme,* 8 (1): 83-98.

Henderson, D. and Horgan, T. (Manuscript) "The Problem of Easy Knowledge and How to Conceive of Epistemic Entitlement".

Hill, K. and Linden, D. (2013) "Hallucinatory Experiences in Non-clinical Populations", in R. Jardri *et al.* (eds.), *The Neuroscience of Hallucination*, pp. 21-41.

Höerl, C. (2001) "The Phenomenology of Episodic Recall", in C. Hoerl and T. McCormack (eds.), *Time and Memory*, Oxford University Press pp. 315-338.
– (1999): 'Memory, Amnesia, and the Past', *Mind & Language 14*, pp. 227-251.
– (2007) "Episodic Memory, Autobiographical Memory, Narrative: On Three Key Notions in Current Approaches to Memory Development", *Philosophical Psychology,* 20 (5): 621-640.

Hopkins, R. (2012) 'What Perky Did Not Show' *Analysis* 72 (3): 431-39.

Howe, (2000)

Huemer, M. (2001) *Skepticism and the Veil of Perception.* New York: Rowman and Littlefield.

– (2007) "Sense-Data", *The Stanford Encyclopedia of Philosophy (Spring 2009 Edition)*, Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/spr2009/entries/sense-data/

Hume. D. (1748/2000) *An Enquiry Concerning Human Understanding* Oxford: Clarendon

Ingle, D. (2005) "Two Kinds of 'Memory Images': Experimental Models for Hallucinations?" *Behavioral and Brain Sciences* 28:6 768

Intraub, H. (1997) "The representation of Visual Scenes." *Trends in the Cognitive Sciences*, 1, 217-221.

– (2002). "Visual scene perception." In Nadel, L. (Ed.) *Encyclopedia of Cognitive Science,* 4, pp. 524 - 527. London: Nature Publishing Group.

Intraub, H., Bender, R and Mangels, J. (1992). "Looking at pictures but remembering scenes." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18:180–91.

Intraub, H., and Richardson, M. (1989). "Wide-angle memories of close-up scenes." *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 179-187.

Jacobson, A. (2005) "Is the brain a memory box?" *Phenomenology and the Cognitive Sciences,* 4: 271-278.

James, W. (1890), *The Principles of Psychology,* 2 volumes, London: Macmillan

Jeshion, R. (2010) "Singular Thought: Acquaintance, Semantic Instrumentalism, and Cognitivism in *New Essays on Singular Thought,* Robin Jeshion, ed. Oxford University Press

Johnson, M. (2007) "Memory systems: A cognitive construct for analysis and synthesis" in *Science of Memory: Concepts*, Eds. H. Roediger, Y. Dudai, and S. Fitzpatrick, Oxford University Press, pp. 353-357

Johnston, M. (2004) "The Obscure Object of Hallucination" *Philosophical Studies* 120 (1-3):113-83.

– (2006) "Better than Mere Knowledge? The Function of Sensory Awareness" in T. Szabó Gendler and J. Hawthorne eds. *Perceptual Experience* Oxford University Press

Jones, S. Guy, A. and Ormrod, J.A. (2003) "A Q-methodological study of hearing voices: A preliminary exploration of voice-hearers' understanding of their experiences" *Psychology and Psychotherapy, Theory Research and Practice,* 76, 189-209.

Jones, S., Fernyhough, C., Larøi, F. (2010) "A Phenomenological Survey of Auditory Verbal Hallucinations in the Hypnagogic and Hypnopompic States" *Phenomenology and Cognitive Science* 9: 213-224

Kaplan, D. (1989a), "Demonstratives," in J. Almog, H. Wettstein and J. Perry (eds) *Themes from Kaplan,* New York, Oxford University Press, pp. 481-563

– (1989b) "Afterthoughts" in Almog, Wettstein and Perry, eds. *Themes from Kaplan,* New York: Oxford University Press, 565-614

King, J. (2007) *The Nature and Structure of Content,* Oxford University Press.

Kolers, P. (1975) "Specificity of Operations in Sentence Recognition", *Cognitive Psychology,* 7: 289-306.

Lackey J. (2007) "Why Memory Really is a Generative Epistemic Source", *Philosophy and Phenomenological Research* 74: 209-219

- (2005) "Memory as a Generative Epistemic Source." *Philosophy and Phenomenological Research* 70: 636-58.
- (2008) *Learning from Words: Testimony as a Source of Knowledge* Oxford: Oxford University Press

Lauria, A. (1987) *Mind of a Mnemonist: a little book about a vast memory,* Harvard University Press.

Lehrer, K. and Richard, J. (1975) "Remembering without Knowing" Grazer Philosophische Studien 1: 121-126

Levine, D. and Finklestein, S. (1982) "Delayed Psychosis After Right Temperoparietal Stroke or Trauma: Relation to Epilepsy" *Neurology* 32: 267-73

Levine, L. (1997) "Reconstructing Memory for Emotions", *Journal of Experimental Psychology,* 126 (2): 165-177.

Lewis, D. (1980) "Veridical Hallucination and Prosthetic Vision", *Australasian Journal of Philosophy*, 58: 239-249.

- (1973) "Causation" *Journal of Philosophy,* 70 556-67
- (1979) "Attitudes *De Dicto* and *De Se*" *Philosophical Review*, 88: 513-543
- (2000) "Causation as Influence" *Journal of Philosophy* 97: 182-97

Leyden, von W. (1961). *Remembering: A philosophical Problem.* London: Duckworth.

Locke, J. (1690), *An Essay Concerning Human Understanding,* in P.H. Nidditch (ed.) 1975 Oxford Clarendon Press.

Loftus, E. & Cahill, L. (2007) "Memory Distortion: From Misinformation to Rich False Memory" in J.S. Nairne (ed.) *The Foundations of Remembering: Essays in Honor of Henry L. Roediger, III"*, pp. 413-427, New York: Psychology Press.

Loftus, E. & Palmer, J. (1974) "Reconstruction of automobile destruction: an example of the interaction between language and memory", *Journal of Verbal Learning and Verbal Behavior,* 13, 585-589.

Loftus, E. (1979) *Eyewitness Testimony.* Cambridge, MA: Harvard University Press. (National Media Award, Distinguished Contribution, 1980). (Reissued with new Preface in 1996). Japanese edition: Seishin Shobo, Tokyo.

- (2005) "Planting misinformation in the human mind: a 30-year investigation of the malleability of memory", *Learning and Memory*, 12: 361-366.
- (1998/1996) "Illusions of Memory", *Proceedings of the American Philosophical Society,* 142(1): 60-73.

Luminet, O. and Curci, A. (2009) "The 9/11 attacks inside and outside the US: Testing four models of flashbulb memory formation across groups and the specific effects of social identity." *Memory,* 17 (7) 742-759.

Mackey, J. (1965) "Causes and Conditions", *American Philosophical Quarterly,* 2/4: 245-55.

Macpherson, F. (2013) “The Philosophy and Psychology of Hallucination: An Introduction” in F. Macpherson and D. Platchais (eds.) *Hallucination* MIT Press

Maine de Biran, P. (1929, 1804) *The Influence of Habit on the Faculty of Thinking,* (republished 1970), Greenwood Press.

Malcolm, N. (1963) *Knowledge and Certainty: Essays and Lectures*, Prentice-Hall, Engelwood Cliffs, N.J.

– (1977) *Memory and Mind* Cornell University Press

Margalit, Avishai (2002), *The Ethics of Memory*, Cambridge, MA: Harvard University Press.

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

Martin, M. (2001) “Out of the Past: Episodic Recall as Retained Acquaintance” in *Time and Memory* eds. Hoerl and McCormack Clarendon Press Oxford

Mast, F. (2005) “Mental Images: Always Present, Never There” *Behavioral and Brain Sciences* 28:6 769-770

Matthen, M. (2010) “Is Memory Preservation?” *Philosophical Studies* 148 (1)

McClelland, J.L., McNaughton, B.L. & O’Reilly, R. C. (1995) “Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory”. *Psychological Review*, 102, 419-457.

McGrath, M. (2007) “Memory and Epistemic Conservatism,” *Synthese* 157: 1–24.

McKay, R. & Dennett, D. (2009) "The evolution of misbelief," *Behavioral and Brain Sciences*, 32, 493-561.

McKinsey, M. (2009) "Thought by Description" *Philosophy and Phenomenological Research* Vol. LXXVIII no 1

McLaughlin, B. (1987) "Why Perception Is Not Singular Reference," in J. Heil *Cause, Mind, and Reality* Dordrecht: Kluwer

Meyers-Shutz and Schwitzgebel 2013

Michaelian, K. (2013) "The information effect: Constructive memory, testimony, and epistemic luck" *Synthese, 190*/12 2429-2456.

– (2011) "Generative Memory," *Philosophical Psychology* 24 (3): 323 – 342

Milner, B. (1962) Les troubles de la mémoire accompagnant des lésions hippocampiques bilatérales. In: *Physiologie de l'Hippocampe*. Paris: Centre National de la Recherche Scientifique, pp. 257-272. (Milner, B.M. and Glickman, S., trans.) Princeton: Van Nostrand, 1965, pp 97-111.

Mitchell, K. & Johnson, M. (2000) "Source Monitoring: attributing mental experiences" in E. Tulving and F. I. M. Craik (Eds.) *The Oxford Handbook of Memory,* Oxford: Oxford University Press, pp. 179-195.

Moon, A. (2013) "Remembering entails Knowing", *Synthese,* 190 (14): 2717-2729.

Moore, G. E. (1953) *Some Main Problems of Philosophy* George Allen & Unwin ltd.

Móró, L. (2010) "Hallucinatory Altered States of Consciousness" *Phenomenology and Cognitive Science* (9) 241-252

Morrison, A. (2001) "The Interpretation of Intrusions in Psychosis: An Integrative Cognitive Approach to Hallucinations and Delusions" *Behavioural and Cognitive Psychotherapy* 29: 257-76

Morrison, A. Frame, L. and Larkin, W. (2003) "Relationships between Trauma and Psychosis: A Review and Integration" *British Journal of Clinical Psyhology* 42: 331-53

Morrison, A. Wells, A. and Northard, S. (2002) "Cognitive and Emotional Predictors of Predisposition to Hallucinations in Non-patients" *British Journal of Clinical Psychology* 41: 259-70

Moscovitch, M. (2007), "Memory: Why the engram is elusive", in *Science of Memory: Concepts*, Eds. H. Roediger, Y. Dudai, and S. Fitzpatrick, Oxford University Press pp. 17-22.

Munsat, S. (1967) *The Concept of Memory.* New York: Random House.

Nadel, L. (2007), "Consolidation: The demise of the fixed trace", in *Science of Memory: Concepts*, Eds. H. Roediger, Y. Dudai, and S. Fitzpatrick, Oxford University Press pp. 177-182.

Naylor, A. (1986) "Remembering without Knowing: Not without Justification" *Philosophical Studies* Vol 49. No 3: 295-311

– (forthcoming) "Belief From the Past" *European Journal of Philosophy* 18 (2)

Neisser, U. (1981) John Dean's memory: a case study", *Cognition,* 9, 1-22.

– (1967) *Cognitive Psychology,* New York: Meredith

Neisser, U. & Harsch, N. (1992) "Phantom flashbulbs: false recollections of hearing the news about Challenger" in E. Winograd, & U. Neisser (Eds.) *Affect and*

accuracy in recall: Studies of 'flashbulb' memories, New York: Cambridge University Press, pp. 9-31.

Neta, R. & Rohrbaugh, G. (2004) "Luminosity and the Safety of Knowledge", *Pacific Philosophical Quarterly,* 85 (4): 396-406.

Neter, J. & Waksberg, J. (1964) "A study of response errors in expenditures data from household interviews," *American Statistical Association Journal,* 59, 18-55.

Nielsen, J. (1958) *Memory and amnesia*, San Lucas Press, Los Angeles.

Nigro, G. & Neisser, U. (1983) "Point of view in personal memories."*Cognitive Psychology,* 15, 467-482.

Noda, S. Mizoguchi, M. and Yamamoto, A. (1993) "Thalamic Experiential Hallucinosis" *Journal of Neurology, Neurosurgery, and Psychiatry* 56: 1224-26

Ost, J. Granhag, P, Udell, J. and Roos af Hjelmsater, E. (2008) "Familiarity breeds distortion: the effects of media exposure on false reports concerning media coverage of the terrorist attacks in London on 7 July 2005", *Memory* 16 (1) 76-85.

Owens, D. (2000) *Reason without Freedom* Routledge

Park, S., Intraub, H., Yi, D-J Widders, D. and Chun, M.M. (2007) "Beyond the edges of a view: Boundary extension in human scene-selective visual cortex." *Neuron* 54:335–42.

Parker, E., Cahill, L. & McGaugh, J. (2006) "A case of unusual autobiographical remembering," *Neurocase,* 12(1) 35-49.

Peacocke, C. (1985) “Imagination, Experience, and Possibility” in J. Foster and H. Robinson (Eds.) *Essays on Berkeley: a Tercentennial Celebration,* Oxford University Press.

Penfield, W. (1959) *The Excitable Cortex in Conscious Man* Liverpool: Liverpool University Press

Penfield, W. and Perot (1963) “The brain’s record of visual and auditory experience: a final summary and discussion” *Brain* 86: 595-696

Perky, C.W. (1910) ‘An Experimental Study of Imagination’ *The American Journal of Psychology*, 21:3 p432-452.

Pezdak, K. (2003) “Event memory and autobiographical emmory for the events of September 11, 2001”, *Applied Cognitive Psychology,* 17: 1033-1045.

Phillips, I. (2012) Book Review of *Memory: A Philosophical Study*, in *Mind*

Pillemer, D. (1984) “Flashbulb memories of the assassination attempt on President Reagan” *Cognition* 16 (1): 63-80.

Pinker, S. & Prince, A. (1988) “On Language and Connectionism: Analysis of a Parallel Distributed Processing Model of Language Acquisition”, *Cognition* 23: 73-193.

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

Plato (1921/360bc) Theaetetus, trans. H.N. Fowler, Loeb Classical Library. London: William Heineman

Pollock, J. (1986) *Contemporary Theories of Knowledge*. Totowa: Rowman and Littlefield.

– (1974) *Knowledge and Justification.* Princeton

Price, H. H. (1952). "Memory." *Philosophical Quarterly* 2 (9):350-355.

Pritchard, D. (2005), *Epistemic Luck,* Oxford: Oxford University Press.

Pryor, J. (2007) "An Epistemic Theory of Acquaintance" available at http://www.jimpryor.net/research/papers/Acquaintance.pdf

Quillian, M. (1968) "Semantic Memory" in M. Minsky (ed.) *Semantic Information Processing,* MIT Press pp 227-270.

Ranalli, C. (2014) "Luck, propositional perception, and the entailment thesis", *Synthese* DOI 10.1007/s11229-013-0323-z.

Recanati, F. (1993) *Direct Reference: From Language to Thought* Oxford: Basil Blackwell

– (2010) "Singular Thought: In Defence of Acquaintance" in *New Essays on Singular Thought* Ed. Robin Jeshion, Oxford University Press

Redelmeier, D. & Kahneman, D. (1996) "Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures", *Pain*, 66 (1): 3-8.

Redelmeier, D., Katz, J., & Kahneman, D. (2003) "Memories of colonoscopy: a randomized trial", *Pain*, 104(1-2): 187-94.

Reisberg, D. (2006), "Memory for emotional episodes: The strength and limits of arousal-based accounts", in B. Uttl, N. Ohta, and A. Siegenthaler (Eds.) *Memory and Emotion: Interdisciplinary Perspectives,* New York, Blackwell.

Reynolds, S. (2002) "Testimony, Knowledge, and Epistemic Goals", *Philosophical Studies,* 110 (2): 139-161.

Richard, M. (1982) "Tense, Propositions, and Meanings", *Philosophical Studies* 41 337-351.

Robinson, J. & Swanson, K. (1993) "Field and observer modes of remembering", *Memory,* 1(3): 169-184.

Roediger, H. (1980) "Memory Metaphors in Cognitive Psychology", *Memory and Cognition,* 8: 231-46.

– (1996) "Memory illusions." *Journal of Memory and Language* 35:76–100.

Roediger, H. Rajaram, S. & Srininvas, K. (1990) "Specifying criteria for postulating memory systems", *Annals of the New York Academy of Sciences,* 608, 572-589.

Roediger, H. Marsh, E. & Lee, S. (2002), "Kinds of memory", in D. L. Medin & H. Pashler (Eds.), *Stevens' Handbook of Experimental Psychology: Vol 2. Memory and Cognitive Processes* (3rd ed. pp. 1-41) New York: John Wiley.

Roediger, H., Dudai, Y. and Fitzpatrick, S (2007) *Science of Memory: concepts* Oxford: Oxford University Press

Russell, B. (1948), *Human Knowledge: Its Scope and Its Limits,* George Allen & Unwin.

– (1912/2004) *The Problems of Philosophy,* Barnes & Noble Publishing.

– (1921/2007) *The Analysis of Mind* BiblioBazaar'

Russell, J. and Hanna, R. (2012) "A Minimalist Approach to the Development of Episodic Memory" *Mind and Language* 27 (1) pp. 29-54

Ryle, G. (1949) *The Concept of Mind,* London: Hutchinson

Sacks, O. (2012) *Hallucinations,* Alfred A. Knopf, a division of Random House, Inc., New York.

– (1970) *The Man Who Mistook His Wife for a Hat and Other Clinical Tales,*

Sainsbury, R. M. (2005) *Reference without Referents* Oxford.

Salmon, N. (1986) *Frege's Puzzle (Second Edition)* (1986). Ridgeview, Atacadero, California.

Sass, L. (1994) *The Paradoxes of Delusion: Wittgenstein, Schreber, and the Schizophrenic Mind.* Ithaca, NY: Cornell University Press

Saunders, J. (1965) "Does all Memory Imply Factual Memory?", *Analysis* 25: 109-115.

Schacter, D., Norman, K., & Koutstaal, W. (1998) "The Cognitive Neuroscience of Constructive Memory", *Annual Review of Psychology,* 49, 289-318.

Schacter, D. (1995), "Memory distortion: History and current status" in D. Schacter et al. (Eds.) *Memory Distortion,* Cambridge, MA: Harvard University Press. pp. 1-43.

– (2007) "Memory: delineating the core", in *Science of Memory: Concepts*, Eds. H. Roediger, Y. Dudai, and S. Fitzpatrick, Oxford University Press pp. 23-28.

– (1990) " Perceptual representation systems and implicit memory: Toward a resolution of the multiple memory systems debate", in A Diamond (Ed.), Development and neural basis of higher cognitive function, *Annals of the New York Academy of Sciences,* 608, 543-571.

– (1996) *Searching for Memory: the brain, the mind, and the past*, New York: Basic Books.

– (2001) *The Seven Sins of Memory* New York: Houghton Mifflin

Schacter, D. & Addis, D. (2007) "The cognitive neuroscience of constructive memory: Remembering the past and imagining the future," *Philosophical Transactions of the Royal Society B,* 362, 773-786.

Schacter, D. Addis, D.R. and Buckner, R. L. (2007) "The Prospective Brain: remembering the past to imagine the future" *Nature Reviews Neuroscience* 8: 657-661.

Schacter, D. and Tulving, E. (1994) "What are the memory systems of 1994?" in Schacter, D. L. and Tulving, E. eds. *Memory. A Volume in the Handbook of perception and Cognition,* Cambridge: MIT Press, 11. 1-38.

Schacter, D., Wagner, A. and Buckner, R. (2000) "Memory Systems of 1999" in E. Tulving and F. Craik (eds) *The Oxford Handbook of Memory,* Oxford University Press.

Schechtman, M. (1994) "The truth about memory," *Philosophical Psychology,* 7 (1), 3-18.

– (2010) "Memory and Identity", *Philosophical Studies* 153 (1): 65-79.

Scheffler, I. (1965) *Conditions of knowledge: An introduction to epistemology and education*, Chicago: Scott, Foresman, and Company.

Schwitzgebel, E. (draft) "Knowing That P Without Believing That P" available at http://www.faculty.ucr.edu/~eschwitz/SchwitzAbs/KB.htm

Senor, T. (1993) "Internalistic Foundationalism and the Justification of Memory Belief" *Synthese* 94, pp 453-476

– (2007) "Preserving Preservationism: A Reply to Lackey", *Philosophy and Phenomenological Research,* 74 (1): 199-208.

Shanton, K. (2011) "Memory, Knowledge and Epistemic Competence" *Review of Philosophy and Psychology*, 2(1): 89-104.

Shoemaker, S. (1970), "Persons and Their Pasts," *American Philosophical Quarterly* 7, pp269-285

Smith, A.D. (2002) *The Problem of Perception,* Harvard University Press

Smith, D. (1983) "Is This a Dagger I See before Me?" *Synthese,* Vol 54 (1) 95-114

Soames, S. (2003) *Philosophical Analysis in the Twentieth Century* Vol. 2, Princeton: Princeton University Press

- (2005) *Reference and Description*, Princeton, Princeton University Press

Sosa, D. (1996) "Getting Acquainted with Perception," *Philosophical Issues* 7: 209-14

- (2007) "Perceptual Friction," *Philosophical Issues 17*: 245-61
- (2011) "Perceptual Knowledge," in S. Bernecker and D. Pritchard eds. *The Routledge Companion to Epistemology* Routledge, Hoboken

Squire, L. (2007) "Memory Systems: A biological concept" in *Science of Memory: Concepts*, Eds. H. Roediger, Y. Dudai, and S. Fitzpatrick, Oxford University Press pp. 339-343

- (2004), 'Memory Systems of the Brain: a brief history and current perspective', *Neurobiology of Learning and Memory*, 82: 171–177.

Squire, L. and Kandel. E. (1999) "Memory: From Mind to Molecules" W.H. Freeman & Co., New York.

Squires, R. (1969) "Memory Unchained", *Philosophical Review*, 78:178-96.

Stanley, J. (2004) "On the Linguistic Basis for Contextualism", *Philosophical Studies,* 119 (1-2) 119-146.

Steup, Matthias, "Epistemology", *The Stanford Encyclopedia of Philosophy (Winter 2012 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2012/entries/epistemology/>.

Stich, S. (1983) *From Folk Psychology to Cognitive Science: The Case Against Belief* Publisher: The MIT Press | 1985-02-20 | ISBN: 0262690926

Suddendorf, T. & Corballis, M. (2007) The evolution of foresight: What is mental time travel and is it unique to humans?" *Behavioral and Brain Sciences,* 30, 299-313.

Surprenant, A. and Neath, I. (2009) *Principles of Memory,* Psychology Press. Taylor and Francis Group, New York.

Sutton (2010a) "Memory", *The Stanford Encyclopedia of Philosophy (Summer 2010 Edition),* Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/sum2010/entries/memory/.

- (2010b), "Observer Perspective and Acentred Memory: some puzzles about point of view in personal memory" *Philosophical Studies 148 (1)*
- (2011), "Memory: a Philosophical Study", *Analysis Reviews,* 72 (1) 181-184).
- (1998), *Philosophy and Memory Traces: Descartes to Connectionism,* Cambridge: Cambridge University Press
- (2004), "Representation, Reduction, and Interdisciplinarity in the Sciences of Memory" in H. Clapin, P. Staines, and P. Slezak (eds) *Representation in Mind* Elsevier
- (2007), "Batting, Habit, and Memory: the embodied mind and the nature of skill" *Sports in Society* 10: 763-786

– (2009), “Remembering” in P. Robbins and M. Aydede (eds) *The Cambridge Handbook of Situated Cognition,* Cambridge: Cambridge University Press, pp. 217-235

Toth, J. and Hunt, R. (1999) “Not one versus many, but zero versus any: Structure and function in the context of the multiple memory-systems debate”, in J.K. Foster and M Jelicic (Eds.), *Memory: System, process, or function?* (pp. 232-272) Oxford University Press.

Traiger, S. (1978) “Some Remarks on Lehrer and Richard’s ‘Remembering without Knowing’” Grazer Philosophische Studien 6: 107-11

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

– (1972) “Episodic and Semantic Memory” in *Organization of Memory,* ed. E. Tulving, W. Donaldson, pp 381-403. New York Academic
– (2002) “Episodic Memory: From Mind to Brain” *Annual Review of Psychology* (53) 1-25
– (2007a) “Are There 256 Different Kinds of Memory?” in J.S. Nairne (ed) *The Foundations of Remembering: Essays in Honor of Henry L. Roediger III* pp. 39-52. New York, NY, Psychology Press xi, 451 pp.
– (2007b) “Coding and representation: Searching for a home in the brain” in *Science of Memory: Concepts*, Eds. H. Roediger, Y. Dudai, and S. Fitzpatrick, Oxford University Press, pp. 65-69.
– (1985) “Memory and Consciousness”, *Canadian Psychology* 26 1-12.
– (1991) “Memory Research is Not a Zero Sum Game”, *American Psychologist, 46,* 43-45.

– (1999) Study of memory: processes and systems. In J. K. Foster & M. Jelicic (Eds.), *Memory: Systems, Process, or Function?* (pp. 11-30). New York, NY: Oxford University Press.
– (2000) "Concepts of memory", in Tulving, E. and Craik, F.I.M., eds. *The Oxford Handbook of Memory*. New York: Oxford University Press, pp. 33–44.
– (2002) "Episodic memory: from mind to brain", *Annual Review of Psychology*, 53 1-25.

Tye, M. (1998) "Externalism and Memory" *Proceedings of the Aristotelian Society, Supplementary Volume*
– (2009) *Consciousness Revisited: Materialism without Phenomenal Concepts* (MIT)
– (forthcoming) "What is the Content of a Hallucinatory Experience?" in *Does Perception have Content?* B. Brogaard (ed) Oxford University Press. Available at: https://webspace.utexas.edu/tyem/www/HallucinatoryExperience.pdf

Unger, P. (1972) "Propositional Verbs and Knowledge", *Journal of Philosophy*, 69: 301-312.

Vignal, J-P., Maillard, L. McGonigal, A. and Chauvel, P. (2007) The dreamy state: hallucinations of autobiographic memory evoked by temporal lobe stimulations and seizures *Brain* 130(1): 88-99 first published online November 30, 2006

Wade, K., Sharman, S., Garry, M., Memon, A., Mazzoni, G., Merckelbach, H. & Loftus, E. (2007), "False Claims about False Memory Research", *Consciousness and Cognition,* 16, 18-28.

Weldon, M.S. (1999) "The memory chop shop: Issues in the search for memory systems", in J.K. Foster and M Jelicic (Eds.), *Memory: Systems, process, or function?* (pp. 162-204) New York: Oxford University Press.

Williams, B. (1973) "Deciding to Believe", in his *Problems of the Self: Philosophical Papers 1956-1972*, Cambridge: Cambridge University Press, 136-151.

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

– (2007) "On Being Justified in One's Head" in *Rationality and the Good* eds. M. Timmons, J. Greco, A. Mele Oxford University Press

Yamada, M. (2011) "Getting It Right By Accident", *Philosophy and Phenomenological Research,* 83 (1), 72-105.

Zemach, E. (1968) "A Definition of Memory", *Mind* 77:526-36.