



The Imagined Past

Citation

Entwistle, Megan. 2024. The Imagined Past. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37379889>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#)

HARVARD UNIVERSITY
Graduate School of Arts and Sciences



DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the
Department of Philosophy
have examined a dissertation entitled
The Imagined Past

presented by Megan Entwistle

candidate for the degree of Doctor of Philosophy and hereby
certify that it is worthy of acceptance.

Signature Selim Berker

Typed name: Prof. Selim Berker

Signature Mark Richard

Typed name: Prof. Mark Richard

Signature Susanna Siegel

Typed name: Prof. Susanna Siegel

Signature _____

Typed name: Prof.

Signature _____

Typed name: Prof.

Date: 8/16/24

The Imagined Past

A dissertation presented
by
Megan Entwistle
to
The Philosophy Department

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Philosophy

Harvard University
Cambridge, Massachusetts

August 2024

© Megan Entwistle

All rights reserved

The Imagined Past

Abstract

What is the difference between remembering the past and imagining it? We can remember things we did from an imagined point of view, like when I recall nervously giving a speech as though I were seated in the audience. We can remember past experiences through a generic representation, like when you picture your childhood walk to school. And we can remember feeling one thing through the emotional lens of another, such as past hopefulness in light of a recent disappointment. My dissertation is about the role of imagination in memory, and what its involvement means for the nature of our knowledge of the past.

Table of Contents

<u>Title Page</u>	<u>i</u>
<u>Copyright</u>	<u>ii</u>
<u>Abstract</u>	<u>iii</u>
<u>Table of Contents</u>	<u>iv</u>
<u>Acknowledgments</u>	<u>v</u>
<u>1. Reference to the Past</u>	<u>1</u>
<u>2. Generic Episodic Memories</u>	<u>33</u>
<u>3. Perspectival Forgetting</u>	<u>66</u>
<u>Bibliography</u>	<u>95</u>

Acknowledgments

This dissertation would not have been possible without the encouragement and contributions of a great number of people. My first thanks are to the members of my committee. I am deeply grateful to my primary advisor, Selim Berker, who has been a generous and responsible mentor to me since my first year of graduate school. Selim encouraged me to trust my philosophical sensibilities in developing my research project, while also offering detailed and careful guidance at the sentence level. He showed me how it's possible to pursue both depth and rigor at the same time. Susanna Siegel inspired me to do philosophy by paying close attention to phenomena from real life. She always helped me find the best way to tell a philosophical story, and she provided generous advice and support for navigating the profession. Mark Richard helped me think through issues at the interface of mind and language that cropped up in every paper I wrote. The confidence he expressed in the project throughout its various life cycles was invaluable in keeping me moving forward.

I owe a debt of gratitude to other faculty members who helped me find an intellectual home in the department. Dick Moran's approach to philosophical questions, in seminars and in conversation, was an early source of inspiration for my research. Ned Hall helped me become a better teacher. And Bernhard Nickel extended a level of mentorship and care that helped rekindle my love for philosophy at the tail end of my degree.

I am immensely grateful to the close friends I made while at Harvard who brought joy to this period of my life. Ege Yumusak, my first friend in the program who was a mile ahead of me from day one, was always ready to look over a draft or strategize about my next step. Erik Knall, with his infinite reserves of energy on and off the ice, empowered me to get the project over the finish line. Samuel Dishaw, whose infectious passion for philosophical conversation made him my first and constant interlocutor, has been there for every step of the journey. He continues to inspire me to believe in the process.

I also owe thanks to colleagues and friends in the wider philosophical community who have helped me grow in conversation, including Simon-Pierre Chevarie-Cossette, Marc-Kevin Daoust, Zachary Gabor, Seth Goldwasser, Johannes Mahr and Malcolm Morano.

To my philosophical friends outside of academia, Sam Milner, Jardena Gertler-Jaffe, and Talia Albert: the thoughts on these pages are for you.

Finally, I want to thank my brother, Joel, and my parents, Susan and Gary, who not only offered endless love and support over the years, but were also there for the bears. My family will recognize just how much of my philosophy is autobiography.

Chapter One

Reference to the Past

Abstract: We come to know the past by accurately remembering events we experienced. Yet whether a memory accurately represents a past event depends on whether the memory successfully refers to the event. How do memories *refer* to their objects? One answer to this question falls neatly out of traditional causal theories of memory: reference is secured by a causal link, sustained through a memory trace. However, recent advances in memory science suggest that remembering is an inherently *constructive* rather than *preservative* process. Motivated by these empirical findings, the new philosophical paradigm treats remembering as a species of constructive imagining. This paper raises a novel challenge for the constructivist paradigm, focussing on its most promising representative: simulation theory. The challenge is to explain how remembering, so understood, could secure the right referential link to past events. I argue that simulation is unable to perform the reference-fixing role previously assigned to memory traces. In dropping the causal condition on remembering, constructive theories have no viable route to explaining memory-based reference to, and therefore knowledge of, the past.

1. Introduction

Episodic memory is thought to be the capacity to represent, from a first-personal perspective, events one previously experienced.¹ I can remember biking to campus yesterday, my best friend's wedding,

¹ The relevant contrast class to episodic memory is *semantic memory*, a distinction originally due to Tulving (1972). The two kinds of memory differ in their objects and modes of presentation: semantic memories are memories of facts, have propositional rather than experiential contents, and are more akin to belief than to perception; episodic memories are memories of events, have experiential contents and phenomenology, and are more akin to perception than belief. This

watching Jaws for the first time. My memories of these occasions relate me to my past in an important way.

Here are two questions we might ask about the nature of this relationship. The first concerns episodic memory as a source of knowledge: how, in virtue of remembering, do we know things about past events? Call this the *Knowledge Question*. The second concerns episodic memory as a form of intentional thought: how, in virtue of remembering, do we refer to past events? Call this the *Reference Question*. These two questions are plausibly related. If I can know how my hockey game went last night because I remember playing in it, my memory must be of last night's game (rather than some other game). Indeed, there is a sense in which the reference question is theoretically prior to the knowledge question. Insofar as my episodic memories supply with me knowledge of the past, they must in general (and in particular instances) provide the right kind of cognitive contact, or referential link, with the past.² But what exactly is the nature of this link, between my episodic representations and the past events they seem to present to me?

On traditional *preservationist* accounts of episodic memory, such as causal theories of memory, the reference question has a straightforward answer. The basic preservationist idea is that the contents of one's previous experience are stored and preserved in what is often called a *memory trace*.³ When a subject remembers some event, she retrieves the contents of her memory trace, ensuring her occurrent representation of the past event has inherited—in a causally appropriate way—the contents of her past experience. The idea of content preservation thus has both an epistemic and referential

essay deals exclusively with episodic memory and episodic remembering; throughout the paper, I will often drop the 'episodic' qualifier and simply say 'memory' or 'remembering'.

² Compare perception: my knowledge about the objects in my immediate environment depends on my perceptual faculties systematically putting me in the right sort of contact with these objects. (What exactly perceptual contact consists in is, of course, a matter of debate: e.g. causal contact, cognitive focus, experiential highlighting, etc.)

³ Prominent examples of causal theories of memory include Martin and Deutscher (1966) and Bernecker (2010). See De Brigard (2014b) for a helpful historical review of the concept of a memory trace and the different explanatory roles it is assigned.

dimension. To illustrate the idea, suppose I have a perceptual experience of a cat at time t_1 which accurately represents the cat before me as fluffy and ginger, such that it could justify a belief at t_1 that the cat is fluffy and ginger. If the representational contents of my perceptual experience are preserved in a memory trace until their retrieval at t_2 , then my recollective experience at t_2 would likewise accurately represent (and justify beliefs about) *that very same cat* as fluffy and ginger.

Memory preservationism, however, has lately fallen out of favor. Burgeoning research on memory errors has shed light on the cognitive processes which underly memory retrieval. For example, the way events are recalled is strongly influenced by contingent features of the subject's retrieval context, such as information she is processing "online" during an attempt to remember.⁴ Brain imagining studies in neuroscience have moreover been taken to reveal that only sparse fragments of information from one's perceptual experience are retained, fragments which are stored not as discrete representations but rather as activation patterns distributed across the subject's sensory cortices.⁵ Together, these empirical findings suggest that the preservationist picture of remembering as the straightforward retrieval of a discretely stored representation of a past experience is, at best, a naïve oversimplification. Taking its place is a new conception of memory, according to which remembering essentially involves an element of *construction*. On constructivist views, remembering subjects piece together representations of an event by drawing on and flexibly recombining different sources of information about it, information which need not causally originate in the subject's past experience of that event.

In this spirit, a number of philosophers have recently argued that remembering subjects heavily recruit the assistance of their faculty of imagination, the idea being that the imagination's productive nature and flexible capacities make it well-suited to explain the new, empirically-backed understanding

⁴ Cf. Roediger and McDermott (1995).

⁵ Cf. De Brigard (2014a).

of memory's constructive underpinnings.⁶ Indeed, some views go so far as to describe memory as a particular species of imagination (namely, imagining directed at one's personal past). Spearheading this program is Kourken Michaelian's *simulationism*, according to which to remember a past event just is to imaginatively reconstruct, or simulate, a representation of that event.⁷

Although simulationism enjoys great empirical support, it faces new philosophical challenges. One challenge is to explain how, if remembering is imagining, episodic memory could supply us with knowledge of the past. The knowledge question has received some treatment by simulation theorists already: Michaelian, for example, has argued at length for the reliability of the episodic construction system that underlies imaginative simulations of the past.⁸

But lying behind the knowledge question is a deeper challenge about reference, which has received comparatively little attention.⁹ The challenge is this. By relinquishing the idea of content preservation through appropriate causation, the simulationist needs to secure a new intentional connection between one's memories, on the one hand, and the objects of those memories, on the other. Evaluations of the simulation accuracy (as well as overall system reliability), which simulation theorists make use of, presuppose there is a fact of the matter as to which event the subject is simulating. The present paper will interrogate the question of what, on an imagination-based view of memory, could ground such a fact about reference.

The paper has two main goals. One is programmatic: I hope to motivate the importance of the reference question for constructivist theories of episodic memory, and to clarify the draws and

⁶ Most notably Michaelian (2016), De Brigard (2014a), and Hopkins (2018).

⁷ Michaelian (2016)

⁸ See Michaelian (2016) for a defense of memory's reliability. The key idea is that the episodic construction system is embedded in a larger metamemory module, which is governed by cognitive monitoring processes that systematically weed out certain kinds of memory errors.

⁹ Simulation theorists are just beginning to consider this problem. See Openshaw & Michaelian (forthcoming).

pitfalls of possible approaches to an answer. The other is argumentative: I make the case that the problem of reference for simulationism is, if not insoluble, very difficult to address. This threatens the viability of the theory.

The paper is organized as follows. In the next section, I lay out, in more detail, the core commitments of simulationism in connection with possible answers they might give to the reference question. I argue that two kinds of approaches to memory reference are available to simulationism. One approach, *descriptivism*, treats the contents of a simulation as descriptive conditions that fix reference to the episode that satisfies it. I motivate different versions of the descriptivist reference principle in section three and argue that no version can deal adequately with the problem of memory errors. The second approach, *intentionalism*, treats the referent of a simulation to be a matter of authorial intent. I explain why intentionalism is a plausible account of imagination-based reference in section four. I then argue that, for these same reasons, intentionalists about memory will be hard pressed to explain how knowledge of the past—and, ultimately, reference to the past—is possible. Because intentionalism eventually gives way to descriptivism, I conclude that we should be pessimistic about the simulation theorist's prospects for delivering a satisfactory answer to the reference question.

2. Simulationism

Michaelian (2016, p. 107) formulates the conditions on episodic memory as follows:

S remembers an episode e just in case

- (1) S now has a representation, R , of e , and

- (2) R is produced by a properly functioning episodic construction system which aims to produce a representation of an episode belonging to S 's personal past.

The role of the episodic construction system is to recruit and recombine different bits of information in order to generate a simulation of the target episode. Not all (or, in principle, any) of the information used in the simulation needs to originate in the subject's previous experience of the target episode, in order for a simulation to accurately represent that episode. It can draw additionally—or instead—on testimony, background beliefs, related perceptual experiences, and so on. The traditional idea that episodic memories are linked to their objects through a non-deviant causal chain, or a memory trace, is dropped entirely from the view.

With the core commitments of simulationism on the table, we can now sharpen the target of our present inquiry. We're looking for a theory of the intentional contents of memory, one that will help the simulation theorist answer the following question.

The Reference Question: In virtue of what does a representation, R , produced by a properly functioning episodic construction system refer to a past episode, e ?

The literature on intentionality of mental states and the varieties of reference is of course vast; the present paper will not pretend to exhaust the possible approaches that might be leveraged for the particular case of memory reference. My orienting suggestion is to conduct the inquiry within the space in which the simulation theorist is already operating. Specifically, I propose we take what simulationism says a memory *is*—or, in simulationist terms, when a subject counts as remembering—to constrain what the theory might say about how a memory refers.

It will be helpful to notice that simulationism replaces virtually all diachronic conditions on memory with synchronic ones. It does not matter, for simulationism, what psychological description we give of the remembering subject *between* the time of the episode and the time of its recollection.¹⁰ The subject need not be in a standing mental state (nor a series of mental states) whose contents match the contents of an initial experience. She need not have stored and retained a discrete representation of the event, and even if she did, she need not draw on that representation to later simulate the event.

What matters instead for ascriptions of remembering are facts about the subject at the time of recollection. The simulationist will want to ascertain, for example, that the subject is recruiting a particular cognitive system (viz. episodic construction) in light of a certain aim or in response to a retrieval cue, that her episodic construction system is not malfunctioning, and that it issues in an occurrent episodic representation.

The simulationist's commitment to synchronic criteria on remembering restricts their range of options when it comes to criterion on referring. Most notably, it excludes diachronic approaches to the reference question, such as causal theories of reference, on which the referential object of one's memory is determined by the causal ancestry of the memory representation.¹¹ There is a fundamental mismatch between a theory of *remembering* which does not require an appropriate causal connection with the remembered event, and a theory of memory *reference* which does require it. The problem, in its general form, is that simulationism countenances memories whose causal ancestry is insufficiently discriminating from a referential point of view. Because memory contents have multiple sources, there

¹⁰ One implication of this is that simulationism does not uphold the intuitive distinction between remembering and *relearning* (where information is forgotten and re-acquired).

¹¹ A causal account of memory reference is operant in Recanati (2007), Soteriou (2018), and of course Martin and Deutscher (1966).

will be multiple causal pathways leading from the memory representation back to distinct past events. In such cases, there would be no fact of the matter about what event the memory refers to.¹²

This leaves us with two places to look for the grounds of memory reference. One is the *output* of the simulation, or the mnemonic representation itself; another is the *process* of simulation, or the cognitive activity the subject is engaged in. Restricting our investigation to these areas would be consistent with the simulation theorist's synchronic conditions on memory.

Let's take the output idea first. How might the representational contents of one's memory relate a subject to her past? Reference for simulationism can't be a function of the causal origins of particular bits of mnemonic content. But it could be a function of how well those contents come together as a depiction of the past episode. I explore a descriptivist theory of reference in section three.

The second option is to think about reference in connection with the simulation process. How might the mental act of remembering relate a subject to her past? Reference for simulationism can't be a matter of a subject's attending to the past via the passive influence of a memory trace. But it could be a matter of a subject's attending to the past via an active, goal-directed undertaking. I explore an intentionalist theory of reference in section four.

¹² As I elaborate in §4.2, the latest cognitive science suggests that, if there are such things as memory traces, the following are likely to be true of them: (1) a single trace is causally shaped by multiple past experiences (Dudai 2012), and (2) a single trace can be (re)used to support simulations of different events (Addis 2020). The simulationist drops any trace requirement from their theory of memory in part because the causal story is so complicated. Adopting a causal theory of reference would thus re-introduce for simulationism many of these same complications. (Cf. Robins 2016b for why *causal constructivism* is a difficult position to maintain.)

3. Descriptivism

Here is a case to motivate the descriptivist approach. I once encountered a grizzly bear close-up during a backcountry hike in the Canadian Rockies. I have a recollection of a stare down between me, frozen in fear, and the bear, haughtily indifferent. What stood out to me most on this occasion were the bear's eyes: pitch black, sunken deep in its head, empty of any remotely anthropomorphic expression. Whenever I think back to the occasion, I picture these eyes. How does my memory pick out that bear? Because the mental image I have of the bear's eyes, were I to describe it, functions as a fitting description of that bear. I remember that bear in virtue of the fit between (i) the features of my memory's contents and (ii) the corresponding properties of its objects. We can formulate the theory as follows:

Descriptivism: R refers to e when the contents of R are satisfied by e .

In more detail, the idea is that the representational contents of a memory serve as descriptive conditions by which reference is fixed. The referential mechanism here is similar to that which governs the use of linguistic referring descriptions, such as definite descriptions: the referent is the object that uniquely fits the description, in virtue of having the properties specified by that description. The difference in the memory case is that the description is supplied through episodic rather than linguistic content.

Descriptivism is a promising account of mnemonic reference for a number of reasons. First, a measure of descriptive fit does appear to drive ordinary judgments of memory reference. Suppose you and I recently went to hear a Wagner opera. You ask me if I can recall how the Parsifal Overture sounds. I think back to the occasion, and hum for you what I remember of the melody. You

immediately recognize the tune I am humming as the version of the “Ride of the Valkyries” featured in *Apocalypse Now*, which I recently watched. Intuitively, your judgment about my memory would be that I am remembering the movie rather than the opera.

Second, descriptivism is a natural proposal for imagination-based reference more generally. John Campbell, for one, proposes a similar explanation for how one’s future-directed thought can pick out objects. He gives an example of a café patron who, while eagerly awaiting the cup of coffee they ordered, imagines the details of their imminent encounter: the warmth of the cup, the steam rising from the surface, the caffeine jolt, etc. Insofar as the customer’s future experience of drinking the coffee in fact fits the description supplied by the contents of his imagining, the customer successfully picks it out in thought.¹³ The descriptivist story of reference-fixing would thus neatly accommodate both future-oriented and past-directed imagining—a feature that should make it particularly attractive to the simulation theorist according to whom remembering is just one instance of a more general capacity for episodic imagining.¹⁴

3.1 The problem of memory error

Simulationism is motivated by decades of research on memory errors and distortions. It is by now a well-documented fact that we are highly susceptible to incorporating misinformation in our recollections of events, or that we misrepresent things in a way that is consistent with our present

¹³ Campbell (2002), p. 177.

¹⁴ Simulationism is committed to the idea of continuity between different modes of episodic imagining: imagining the actual past, imagining the counterfactual past, and imagining the future. The continuist thesis is that these three forms of imagining are only differentiated according to their temporal and modal dimensions but are, at root, the same psychological kind. For recent defenses of “discontinuism,” see Robins (2020) and Langland-Hassan (2021).

beliefs or self-image.¹⁵ I probably remember the bear being larger than it was or represent myself standing closer to it than I did. The color of the bear's eyes or fur in my memory image have probably shifted, since the time of my initial perception, to converge on the look of the various grizzlies I have seen in pictures and videos since.¹⁶ Do I nevertheless remember the encounter, that is, have a memory that counts as being *of* it, despite getting many aspects of the past episode wrong?

The simulation theorist will want to say yes. As they point out, our memories in general suffer (and sometimes benefit) from various content distortions. The difference between remembering accurately and remembering inaccurately does not, on their view, mark a difference in psychological kind.¹⁷ Misremembering is still remembering, just less successfully. And in order to countenance the notion of inaccurate memories, we need to at least in principle be able to specify which past episode the memory is an inaccurate representation *of*. The simulationist's theory of reference should thus be able to accommodate cases of memory errors.¹⁸

But now descriptivism is beginning to look like the wrong tool for the job. Suppose my memory of the bear does include representational contents that are not accurate. If we insist that my memory is nevertheless about that bear encounter, then we have a case of reference without content satisfaction, which violates the descriptivist principle. Call this the *problem of memory error*.

¹⁵ See Loftus (1996; 2005) on suggestibility and eyewitness testimony; see Schacter (2021) for a discussion of consistency bias.

¹⁶ Categorical bias is a well-documented phenomenon for visual memory, in particular for working memory of colors. For example, we tend to recall non-standard shades of blue as prototypical blue. See Zhou et al. (2022); Bae et al. (2015).

¹⁷ See De Brigard (2014a) for an example of a simulationist who emphasizes the importance of distinguishing between questions about what counts as a memory, on the one hand, and questions about a memory's epistemic credentials, on the other.

¹⁸ Simulationism is at odds with memory disjunctivism in this respect. Those who treat remembering as factive (e.g. Bernecker 2010) will tend to deny there is a unified mental state of *memory* which encompasses inaccurate representations of the past. Debus (2008), for instance, defends a version of memory relationism from which it follows that "recollective memories and memory-hallucinations are occurrences of very different kinds" (414). Memory disjunctivists will face no similar theoretical pressure to say that false memories refer.

In what follows, I will go through two ways the descriptivist might respond to the problem of memory error. The first strategy (§3.2) is to accept the result that false memories do not refer while denying that this result would commit the simulationist to an untenable form of memory disjunctivism. The second strategy (§3.3) is to amend the principle of descriptive fit in a way that makes it more permissive to error. I will argue that neither attempt to meet the challenge leaves the simulation theorist with a viable account of description-based reference to the past, and that they would be better off abandoning descriptivism altogether.

3.2 The existential defense

Some memories don't refer. So what? If there is no e that satisfies the contents of R in my memory, then I don't succeed in remembering anything. If our theoretical task is to specify the connection that must obtain between a memory representation and a past event in order for me to have knowledge of the latter on the basis of the former, then it looks like the descriptivist will get things right for cases of memory errors: no reference, and (therefore) no knowledge.

Moreover, the defense continues, no reference-verdicts need not threaten a disjunctive analysis of memories qua psychological kind. That's because the descriptivist can give a unified analysis of the contents of memories—across variation in what we've been calling 'accuracy'—that treats descriptions like existentially quantified propositions. For example, the proposed content of my bear memory is: <There was an event in the past involving a grizzly with such-and-such properties, someone named 'M.E.', on a trail with such-and-such properties... *etc.*> Because there was no such event, the memory is false, and doesn't ground knowledge. If there were such an event, it would be true, and could ground knowledge. In either case, the contents are perfectly meaningful. The existential defense thus appears

to be sufficiently aligned with the simulationist's insistence that the difference between remembering and misremembering is epistemic rather than psychological.

However, the existential analysis of descriptions renders the thoughts or utterances in which these descriptions are couched inherently *general* by ranging over the past as such. For our purposes, what this means is that when I have a false memory, I am not misremembering any one thing—not the actual bear incident, nor any other past event. In recalling that an event-description obtained, which did not obtain, I am as it were misdescribing the past, full stop.

This picture of how memories represent the past strikes me as deeply dissonant with ordinary memory discourse, particularly in contexts of misremembering. Here is something we regularly do: we disagree over “what happened”. For example, my parents were also present at the bear encounter, and we like to argue about how exactly the events unfolded and offer corrections to each other's memories (e.g. “No, you're misremembering: I was the one holding the bear spray”). The best explanation of what makes this kind of discourse intelligible is that we are each able to refer to one and the same occasion, despite disagreeing over its details. In these kinds of cases, charges of misremembering are said in relation to a target episode.¹⁹

Moreover, there is pressure to say that I surely stand in *some* mnemonic relation to the bear. For one thing, I am better positioned to recall it than, say, a marmot further down the trail that didn't show up in my past perceptual experience at all. The present view, on which false memories simply lack intentional objects, is forced to treat the bear and the marmot as on a par. And yet, my memory representation fits the bear much more closely than it does the marmot. This is not a coincidence; rather, the fact that I was previously aware of the bear but not of the marmot explains—at least

¹⁹ I've added the qualifier “in these kinds of cases” to signal that some forms of memory disagreements might have a different structure. Sometimes we combine the core elements of two distinct occasions in a single memory representation. When it's evident that one has done so, the natural mode of objection is not “You're misremembering *x*” or “You're misremembering *y*”, but rather, “You're mixing up *x* and *y*; those were separate occasions.”

partially—why my present memory has <bear> but not <marmot> contents. Indeed, part of the motivation for memory descriptivism in the first place was the idea that we largely take our cues of memory reference from descriptions of memory contents. Surely, this method can succeed even when the contents include bits of false information about the referent. Rather than deny reference for memory errors, the descriptivist would be better off amending their condition of descriptive fit to something more permissive.

3.3 The permissive defense

Suppose the descriptivist proposes a new principle. Call it best-fit descriptivism, or descriptivism_{BF}:

Descriptivism_{BF}: *R* refers to the episode *e* whose properties best satisfy *R*'s contents

The amended principle will now be able to accommodate cases of memories that refer despite some representational inaccuracy. Because my memory contents include, roughly, <bear encountered while hiking>, and because my past bear encounter while hiking fits that description more than any other episode—such as the time I saw a black bear at a campsite or the time I hiked that same trail without encountering wildlife—it successfully picks out the first of the three alternatives.

'Best fit' is a comparative notion. This means that the actual degree of fit between *R* and *e* will depend on the range of relevant alternatives to *e*. If I have encountered bears while hiking on numerous occasions, my memory would need to represent its object more closely than it would if there were just one bear encounter in my past. It would need, for instance, to represent a distinguishing feature of the bear's behavior, whether there were cubs present, whether it was wildflower season, etc.

If, by contrast, there were no such relevant alternatives to e , I could get these features wrong or exclude them altogether yet still successfully refer.

These facts seem to be reflected in ordinary discourse surrounding memory reports: the individuation question (“Which time are you thinking of?”) is appropriate in contexts where multiple options are salient, and an informative response will be one that supplies a differentiating detail (“The time that such-and-such happened”).

On best-fit descriptivism, reference is a function of how well the contents of the memory represent the remembered episode as compared to any other relevant alternative. Total accuracy is not required, but just enough accuracy to (i) perform the individuating work and (ii) render it more plausible that one is remembering the episode that best fits the description than no episode at all.²⁰

One component of descriptivism_{BF} requires elaboration, and that’s the idea of a relevant alternative. What range of episodes count as possible targets of one’s memory? A first pass response is: anything that the subject is in a position to remember. You are not in a position to remember, for instance, events you had no previous experience of, regardless of how well any mental representation of yours fits some such event. But we need to tread carefully here. We are considering descriptivism for simulationism in particular. Simulationism, not descriptivism, supplies the constraints on what counts as a memory:

- (1) S now has a representation R of e , and
- (2) R is produced by a properly functioning episodic construction system which aims to produce a representation of an episode belonging to S’s personal past

²⁰ Here’s the rationale for including the second condition. The descriptivist may want to posit a lower bound to fit, i.e. some minimal level of descriptive accuracy, to rule out cases in which the best fitting episode is nevertheless too poor of a fit to be a plausible referent. For example, if I never had seen a bear hiking and my family had been gaslighting me all this time, we would perhaps not want to say that my memory is of the next closest thing, such as the campsite bear.

Descriptivism_{BF} merely specifies how a memory—when it counts as such—secures its representational objects. Therefore, the relevant question is: what, on the simulationism–descriptivism_{BF} combination, is the subject in a position to remember? And the answer has to be: whatever episodes from the subject’s personal past the contents of her memory could descriptively fit. We’re moving in a tight theoretical circle.

A problem looms. The problem is that the range of relevant alternatives for any e will include the kinds of cases which—from a referential point of view—we should want to rule out as possible referents. To illustrate, consider the following scenario.

SNOWBOARDING. I have tried snowboarding three times. The first time (e_1) the mountain conditions were good, but I crashed. I attempted it again the next year, only this time (e_2) the conditions were icy, so I crashed again, hit my head, and forgot about the incident shortly after it happened. The third time (e_3) I got lucky, managing to stay on my feet despite the conditions being icy. Years later, I get asked by a friend whether I’ve ever experienced a snowboard wipe-out. Cued with that question, I reconstruct a representation of crashing on icy slopes. I answer my friend in the affirmative, describing my fall.

What episode have I described in SNOWBOARDING? The contents of my memory include the various sensations of falling hard (which fit e_1 and e_2) plus the look and feel of the icy slopes (which fit e_2 and e_3). My representation of the crash is therefore better satisfied by e_2 than it is by either e_1 or e_3 . According to best-fit descriptivism, my memory refers to the second episode.

This should strike the reader as a bad result for the descriptivist. Although e_2 best fits my memory, I am surely not in a position to represent it in thought in a way that would sanction an

ascription of a *de re* attitude to me. That is, we wouldn't want to say that I believe *of* e_2 that I wiped out on the ice.²¹ If a third party to the conversation happened to be present for both e_1 and e_2 , presumably they would correct me, reminding me that e_1 wasn't icy—their assumption here, of course, being that there's no way I could have meant to describe e_2 . And upon receiving this information, the rational thing for me to do would be jettison the representation of icy conditions from my memory.²²

To rescue descriptivism_{BF}, it's therefore tempting to exclude e_2 from the range of relevant alternatives on the grounds that I'm not in a position to remember e_2 . But insofar as the descriptivist takes on the commitments of simulationism, this move is not available. Per simulationism, the fact that one doesn't store and retain a representation of an event does not preclude one from remembering it via imaginative construction, for the reconstructed contents of a memory need not causally originate in the past episode it represents.

It is also tempting to object that the representation I form of the crash should not count as a memory because the concussion I sustained at e_2 adversely impacted my memory system (i.e., the case violates the second condition of simulationism). But it's hard to see how this line could be sustained. The simulation theorist is concerned with how the episodic construction system is functioning at the time of recollection. And we can explain how I combine <crash> with <icy> contents in ways that are par for the course for any healthy memory system. Perhaps I have formed a strong association over the years between ice and slipping. Perhaps I believe that snowboard crashes are more likely to

²¹ Prominent examples in the reference literature support a similar analysis. Kvat's (1989) hallucinating subject's representation of a burglar in the basement does not pick out the actual burglar in the basement (whom the subject neither sees nor hears), even if his representation is a perfect match of that individual. Donnellan's (1996) speaker at the party who asks, "Who is the man drinking a martini?" refers to the person in the speaker's visual field who is holding a water-filled martini glass, rather than someone in the next room who is in fact drinking a martini. In both of these cases, the intuition that the representation does not pick out the best fitting object seems to be tracking the conspicuous absence of an appropriate causal connection between the contents of the subject's mental state and the best fitting object.

²² See Kvat (1989) for a defense of the following idea: the ways that a subject decides to update her representation, upon receiving corrective inputs from other people, is evidence for which descriptive elements (e.g. <icy> or <wipe-out>) play the referential anchoring role for her representation.

happen in icy conditions. Perhaps I am unwittingly incorporating the testimony of a friend who witnessed e_2 . Conceptual associations, standing beliefs about event types, and testimony can all play a guiding role in any ordinary memory formation.

In fact, SNOWBOARDING features exactly the kind of memories that simulationists are willing to take on board.²³ They may for this reason accept the descriptivist verdict here. But doing so would come at the cost of being able to illuminate the epistemic role of memory. If remembering is a matter of descriptive accuracy all the way down, we lose any clear handle on the idea that memory is a distinctive source of knowledge of the past—that is, the kind of mental state whose obtaining in a subject explains why their beliefs about the past count as knowledge. Memory is not *distinctive* if there is no principled difference between the two kinds of response to the question, “How do you know e happened?”: (1) “Because I remember e happening,” and (2) “Because I was told.” And memory is not an epistemic *source* if the fact that your knowledge of e happening explains your remembering e , rather than the other way around.

Happily, the simulation theorist need not be so revisionist. The intuition in SNOWBOARDING is that my memory refers to e_1 . Why? Because surely that’s what I intended to recall; I just made a mistake during reconstruction. The better option, it seems, is to explain reference in connection with the subject’s intentions.

²³ Of a similar case to e_2 , Michaelian says: “The correspondence between the apparent memory and the accident is not a coincidence, as properly functioning episodic construction systems are designed to draw on testimonial information in order to simulate past events” (2016, p. 110). His discussion here concerns what counts as a memory, not how memories secure their representational objects. It is thus open to him to say of SNOWBOARDING that while I could in principle count as remembering e_2 , the more intuitive verdict about reference is that I am misremembering e_1 .

4. Intentionalism

What we have learned from the problem of memory error for descriptivism is this: accuracy is not theoretically prior to reference. Although the recognition that someone's memory seems to describe some episode better than another ordinarily guides our judgment that the person is remembering that episode (rather than the other), descriptive fit does not itself determine that fact.

In rejecting a causal condition on remembering, simulation theorists already distance themselves from the idea that memories reveal their origins through their representational contents.²⁴ It is a short step from here to rejecting the idea that memories reveal their referential objects through their representational contents. The simulationist ought to look beyond the contents of a simulation for an answer to the reference question.

It will be helpful to locate a new starting point for the simulation theorist, in light of the difficulties facing descriptivism. Here is one: there is more to memory than descriptive contents. My memory of yesterday's commute to work, and my imagining tomorrow's commute to work, plausibly share the same mental imagery—but I can nevertheless distinguish what these representations signify in connection with the attitudes that are bound up in my acts of episodic construction. For instance, I treat my remembrance as representing something actual, and in the past, whereas I treat my imagining as representing something possible (or probable), and in the future.²⁵

²⁴ That is, simulation theorists reject Tulving's (1972) "contextual wealth" hypothesis for episodic memories. Tulving proposed the following distinction between semantic and episodic memories: whereas one can remember a fact while forgetting where one learned that fact, one cannot in the same way remember a past experience while forgetting *its* origins (since one's representation of the past experience will necessarily represent the source of one's memory of it, i.e. the past experience itself).

²⁵ It is important to note that the simulationist would largely agree with the thrust of this paragraph. Although they treat remembering the past and imagining the future as two species of episodic imagining, they claim there is nevertheless a difference (and a way for the subject to tell the difference) between imagining the past and imagining the future. Michaelian argues, for example, that "an episodic construction system that did not differentiate between past and future, actual and counterfactual, likely could not have evolved" (2016, p. 106). How exactly this differentiation happens—i.e. what makes it the case that some mental representation picks out a past episode rather than a future one—is related to our present concern about reference.

How do I do this? One influential answer emphasizes that we exercise a distinctive form of first-personal, rational authority over certain kinds of cognitive activities, such as intending to φ or believing that p . This is the authority to determine, without the need to observe oneself or infer from one's behavior, what one is doing.²⁶ This kind of authority extends to imagining. For example, when asked what one is imagining, a person does not ordinarily take this as a request to trace the origins of their imagery in past experience. Nor does a person usually need to turn their attention inward to read off the identity of the objects featured in their imagining, in the way that one would inspect the visual contents of a blurry photograph to make identifications of people and places. Rather, a person can immediately report on the contents of one's imagination insofar as one is the author.²⁷

We can now distill these ideas into a thesis about imaginative reference: the contents of a subject's imagining are about specific objects or events insofar as the subject *intends* to represent those objects or events by imagining. This thesis is friendly to the simulation theorist for two reasons. First, it is motivated by ordinary observations about the practice and nature of imagining (e.g., imaginers do not discover but rather decide what they imagine), and simulationism is at root an imagination-based theory of episodic memory. Second, the seeds of the idea that reference is a function of the remembering subject's intentions are already found in Michaelian's (2016) exposition of the theory. The second condition on S 's remembering some episode e invokes the idea of a properly functioning episodic construction system which "aims" its operations at a representation of the actual past. One way to interpret what it is for the memory system to have an aim is to understand S as initiating an act of episodic construction with a particular target in mind. As Michaelian suggests in passing, "the

²⁶ I am drawing here on the notion of first-personal authority defended by Richard Moran (2001), with roots in Anscombe (1957). If I decide to φ , and proceed to act in accordance with my intention, I have practical self-knowledge that φ -ing is what I am doing.

²⁷ This idea that imaginative contents are fixed by authorial intent—and thus immediately knowable by their authors—has been defended by many philosophers. For famous examples, see Sartre's (1940) theory of imagining as an act of positional consciousness and Wittgenstein's (1958) remarks about why imaginers can't be wrong about the objects of their imaginings.

obvious mechanism to which we might appeal is intention—either the intention of the subject himself or the ‘intention’ of his episodic construction system” (112).²⁸

The intentionalist proposal is this:

Intentionalism: R refers to e when R is the result of an imaginative process undertaken with the intention of representing e .

Here is the plan for discussion of intentionalism. First (§4.1) I motivate an initial epistemological challenge for the simulationist’s adoption of the intentionalist thesis. Then (§4.2) I develop a response on behalf of the simulationist, combining a strategy from the epistemology of imagination literature with constructivist principles of memory. I ultimately argue (§4.3) that this response reveals a dilemma for simulationism, between explaining the possibility of knowledge and explaining the possibility of reference.

4.1 The epistemic power challenge

Although intentionalism is in the first instance a theory of reference, its viability depends on how well it pairs with other central requirements on a theory of memory. Any story about how remembering subjects make cognitive contact with the past should be offered with an eye towards the question of how remembering subjects have knowledge of the past (i.e. what makes remembering well possible at all). Although, as I’ve argued, the Reference Question is in some sense prior to the Knowledge Question, it is not completely theoretically independent. This is because, as I’ll now go on to argue,

²⁸ Michaelian elaborates that we can think of a system’s “intention” in terms of its responses to retrieval cues, such as environmental triggers or person-level prompts.

certain features of content determination have downstream effects on the epistemic status of these contents.

Take the case of imagination. The very same features which make intention a plausible account of the contents of imagination—viz. imagining subjects can't be *wrong* about what it is they imagine insofar as they exercise some kind of *voluntary control* over their imaginings—make it difficult to explain how the contents of imagination could ever serve an epistemic function.

This idea is at work in various challenges to the idea of imagination-based knowledge. How can you learn something through imagining if what you imagine is up to you?²⁹ If imagining is merely a transfer of contents from intention to image, isn't it epistemically inert rather than epistemically productive?³⁰

Such questions are often motivated by a contrast between imagining and other cognitive processes like *perceiving* or *reasoning* that are paradigmatically justification-conferring. Although I can intentionally direct my conscious attention towards some things in my perceptual environment while ignoring others (as when I read a book) or decide which perceptual environment to place myself in (as when I walk to the library), I cannot *decide* what to see and thereby see it, in the same way I can decide what to imagine and thereby imagine it. Similarly, although I can intentionally open an inquiry, choose to gather evidence, or steel myself to work through a complicated math proof, I cannot choose to believe whatever I please at will—rather, the contents of my beliefs are responsive to the reasons I have and issue automatically from the conclusions of my deliberations.

The disanalogy between perceiving and reasoning, on the one hand, and imagining, on the other, makes it difficult to see how the intentional aspect of imagining could leave room for an epistemic aspect. This difficulty is not normally extended to memory. On a preservationist conception

²⁹ Magdalena Balcerak Jackson (2018) calls this the “up-to-us challenge” for imagination.

³⁰ Cf. Peter Langland-Hassan (2016).

of remembering, memory falls squarely on the side of perception and reasoning when it comes to our agential control over its contents. But according to simulationism, remembering is a species of imagining. If the contents of one's imagining are up to the imaginer, then by the same token so too are the contents of one's memory. How then could one's memories be beholden to the past in an epistemically appropriate way? Call this the *epistemic power* challenge for simulationism.

4.2 Imagining under constraints

Several extant views in the epistemology of imagination literature make a case for the imagination's epistemic powers.³¹ Some argue that our imaginations are “constrained” in epistemically appropriate ways by, e.g., the target inquiry, or a structural feature of the imaginative process. These constraints help to ensure that agents do not—at least in certain contexts—have too much top-down control over their imaginings. And in the absence of top-down control, exercises of the imagination can be truth-tracking or justification-conferring after all.

For example, Williamson (2007, 2016) defends what he calls the *cognitive view* of imagining, according to which the imagination plays a distinctively instructive (as opposed to creative) role in a variety of cognitive tasks. In particular, we recruit the imagination to think through the consequences of hypothetical situations. A large portion of our knowledge of counterfactuals (i.e. knowledge of the form: if p were true, q would be true), Williamson argues, are based on imaginative simulations of a world in which some proposition, p , is held fixed and—as the relevant details of that world get filled out—another proposition, q , is revealed to follow. For example: if I imagine attempting to jump across a rapidly moving stream, and—after multiple simulations of myself executing that action—imagine

³¹ Many of these views are domain specific; for example, imagination has been defended a source of modal knowledge (cf. Kung 2010; Balcerak Jackson 2013) and as a way of understanding another person's mental states (cf. Bailey 2022). Because these views are not obviously extendable to imagination-based knowledge of the past, I won't discuss them here.

myself landing successfully on the other side, then I can safely believe that I would in fact clear the stream (q) were I to jump (p).

Important for Williamson's view is that imaginative simulations proceed *involuntarily*. It may be up to the subject when to initiate the episode and how to fix the antecedent content of the imagined scenario; but the way in which the conditional unfolds after that point is not something over which she has agent-level control. Instead, one's imaginative simulations are guided by 'offline' perceptual judgments or sensory input prediction mechanisms, are constrained by norms of reasoning, and draw on one's background knowledge. They tend on the whole, therefore, towards minimal alteration of the actual world as the agent knows it. This is why, Williamson argues, the contents of one's cognitive imaginings are by and large truth-tracking and qualify as a reliable source of counterfactual knowledge.

Williamson's model of imagining involves three stages. First, a target is fixed by one's voluntarily imagining an antecedent to a counterfactual conditional ("if p were to obtain..."). Then, the imagination simulates the unfolding of that hypothetical scenario, filling in content not just in any way the author pleases, but through offline contact with other mental modules such as perception or reasoning. Finally, in the event that the simulation process terminates in a stable imaginative state, the subject is justified in using its contents as evidence for a counterfactual belief. We can call this model of instructive uses of imagining the *constrained imagining* model.

How might the constrained imagining model be helpful to simulationism? The appropriate target of a remembrance is, of course, the actual past rather than the counterfactual past. The target is fixed by a retrieval cue. The cue might be supplied by a question (e.g. What restaurant did we eat at last week?), the purpose of a conscious inquiry (e.g. mentally tracing one's steps to search for lost keys), or an unbidden image or sensation (e.g. a feeling you've forgotten something important after leaving for the airport).

Once the target is fixed by a retrieval cue and the simulation process is initiated, what are the relevant constraints on reconstructing one's past? Here we can look to the memory literature itself for guidance on this question. Two varieties of memory constructivism—connectionism and episodic hypothetical thinking (EHT)³²—posit what are called *distributed memory traces*.

According to connectionist accounts of memory, the episodic information from a subject's experience of a certain event gets stored in a semantic network of "nodes" (corresponding to basic information units, such as concepts or event features). As particular episodes are compiled over the course of a person's lifetime, activation patterns develop in the general network of information which reflect regular associations between certain features of experience. The connections between nodes grow stronger the more those nodes are "co-activated" in the subject's experiences: <birthday> and <cake>, for example, will exhibit strong connectivity. When a person undertakes to recollect any particular occasion, she will use a retrieval cue to fix her guiding node and reconstruct her memory through connections to other nodes that are most frequently co-active with the cue or most recently co-active.³³

A key idea here is that the network's given activation pattern at any time of recollection is a major determining factor in how the subject reconstructs her past. There are a couple of ways to cash out the nature of the distributed memory trace's influence. The influence could be understood negatively: it might *constrain* the subject's reconstruction of an event, for example by limiting her building blocks to a subset of the total available nodes. Or it could be understood more positively: it might *guide* the subject's reconstruction, for example by making connections between certain nodes

³² Connectionism is a theory of the constructive mechanisms that underpin episodic memory generally. It is compatible with, but need not entail, the simulationist's claim that the faculty of *imagination* performs the relevant constructive work. By contrast, De Brigard's (2014a) EHT theory of memory is itself a variety of simulationism.

³³ The first articulation of connectionist memory traces is in Sutton (1998). "[Traces] employ superpositional storage: memories are blended, not laid down independently once and for all, and are reconstructed rather than reproduced." Recent proponents of a similar view of distributed traces include Bernecker (2010) and Michaelian (2011).

more salient to the subject and thus more easily retrievable. Both explanations are compatible with the idea that the reconstruction process is not entirely under the subject's voluntary control.

Once the worry about epistemic downgrade is avoided, there is further room to argue that reconstructed memories are generally truth-tracking when it comes to the past. A second key idea from connectionism is that memory traces preserve event *patterns* rather than *particular* events.³⁴ This explains our capacity to remember general features of event-types while sometimes getting specific details wrong. On the whole, then, our reconstructions of the past will tend to reflect what is *likely* to have occurred, given one's total background experiences. And if reconstructions are pattern-responsive in this way, an argument could be made for their being reality-oriented (in the same way that Williamson's imagination is reality-oriented by recruiting one's background knowledge and predictive mechanisms, to develop the most likely simulation of a counterfactual scenario).

Felipe De Brigard (2014a) argues for a similar conclusion. His version of simulationism, EHT, explicitly weds remembering to counterfactual imagining by arguing that the same reconstruction mechanisms involved in remembering also guide one's thinking about what *could have been* the case and what *will be* the case.

On De Brigard's conception of distributed memory traces, the contents of a subject's perceptual experience are stored in a *gappy* and *disaggregated* fashion. Since traces of past events are neither complete representations nor discretely bound together, it follows that a subject's recollective experience of a past event—which, to her, *is* a discrete representation—requires both the consolidation of information fragments as well as added content. In other words, there is the need for reconstruction in memory. Reconstruction is itself probabilistically constrained by a subject's

³⁴ Robins (2016a)

schematic knowledge (e.g., that feature x typically appears in event type y) as well as frequency of the subject's prior encounters with objects or events of the relevant type.³⁵

Combining the insights of both connectionism and EHT, finally, we can sketch a constrained imagining model of episodic memory. The intended target is fixed by a retrieval cue. The constraints on reconstruction are supplied by distributed memory traces associated with that cue (patterns of activation encoded in a neural network) as well as a subject's schematic knowledge and probabilistic priors. Together, these two aspects explain how it would be possible for a subject to both (i) intend to construct a representation of some episode and (ii) do so accurately.

4.3 The intentionalist's dilemma

There is a lingering worry, however, that the model developed in §4.2 does not adequately respond to the epistemic power challenge. A simple observation gets the worry off the ground: what actually did transpire in any given situation can diverge in any number of respects from what might be expected to happen in a situation of that type. For example, suppose I typically score whenever I play hockey, only last night I didn't. I am not thereby unable to remember that uniquely frustrating game or that particular detail about it. A view on which I remember by piecing together nodes of information that fit the pattern for <hockey game> can't seem to account for this. The explanatory power of distributed traces and probabilistic reconstruction mechanisms thus seems limited to the subset of past events that are more predicable than memorable, so to speak.³⁶

³⁵ De Brigard (2014a); De Brigard et al. (2013)

³⁶ Robins (2016b) gives a related argument against distributed trace views of memory. She draws a different conclusion: that distributed trace theories can't explain how one could count as remembering one pattern-conforming episode rather than any other, since the episodic construction processes involved will be virtually the same.

One way to shrink the gap between optimal and actual reconstructions of the past would be to tighten the link between the target episode and the constraints on imaginative reconstruction. For example, if in the course of reconstructing my hockey game I also drew on, say, patterns of experiences of feeling disappointed, background beliefs about causes for disappointment related to hockey games, and so on, I would presumably be able to imagine something closer to the actual event. This suggestion, unfortunately, solves one problem by generating another: now, the theorist is on the hook for specifying a way of fixing the target of past-directed imagining in a way that (i) doesn't require the subject to possess an antecedent conception of that episode, yet (ii) nevertheless would help the remembering subject to select the epistemically appropriate inputs to her simulation. In other words, the problem is that the ability to recruit the right bits of schematic knowledge or frequencies to guide one's reconstruction seems to depend on the subject already knowing what a successful reconstruction would look like (or, at the subpersonal level, depend on the system having stored a discrete representation). But the epistemic power of imagining is supposed to explain how we could have knowledge of the past, not the other way around.

Let's take stock. We are considering an intentional theory of mnemonic reference according to which the reference of one's simulation is fixed through authorial intent. The idea of authorial intent raises an immediate concern about the possibility of mnemonic knowledge: for if my intending to imagine makes it so that I imagine it, how could my act of imagining ever constitute evidence of (the possibility of, or actuality of) what the contents of my imagination present? A potential response to that challenge, leveraged from the nearby epistemology of imagination literature, was to restrict the scope of authorial intent to the initiation phase of the simulation process, so that the real epistemic work of imagining is something carried out under appropriate constraints. In the course of exploring how the target-constraints-output model of imagining could apply to constructive remembering (specifically, simulationism), a new version of the initial problem emerged. The problem was that

imagining the past *successfully* under constraints requires a tight enough connection between the intended target and the constraints supplied by that target, where the effect of tightening that connection shifts the locus of epistemic power back to the stage of target-fixing.

The intentionalist theory of reference thus presents a dilemma for the simulation theorist: either (1) explain reference to the past in a circular way or (2) concede that imagining the past is not conducive to knowledge. By my lights, (1) is more worrisome than (2).³⁷ If an intentional theory of memory reference is to be minimally adequate, it should not build in the requirement that subjects remember past episodes *before* they simulate them. The best version of the intentionalist view of memory, then, is one where all three conditions are satisfied:

- (i) S intends to remember some particular episode, e .
- (ii) In forming this intention, S does not draw on any substantive information S has retained about e .
- (iii) As a result of intending to remember e , S 's imaginative simulation (whether accurate or not) refers to e .

Suppose, for the sake of argument, that the simulationist concedes on the epistemic front, accepting that imaginative reconstructions of the past will often miss the mark. To evaluate the intentionalist proposal qua theory of reference, we now need to ask whether conditions (i) through (iii) could all be met in both successful and unsuccessful instances of reconstruction.

One way to model intending to remember, such that conditions (i) and (ii) are met, is by way of minimal descriptions. Suppose my friend asks me what I did last Tuesday. I can use the description

³⁷ Constructive theories of episodic memory are first and foremost theories of the *process* of remembering. It would be a worse result for simulationism to be forced to countenance discretely stored representations to explain the possibility of memory, than for simulationism to be unable to explain how constructive processes reliably produce accurate memories.

“last Tuesday” to fix the target of my recollective inquiry without requiring any antecedent conception of what I did last Tuesday. I now begin to reconstruct the events of the day. Here’s one way the simulation process might proceed.

LAST TUESDAY. Using “last Tuesday” as a retrieval cue, I recruit my standing beliefs about my weekly semester schedule, viz. that Tuesdays are the day I usually teach. My Tuesday-teaching association generates a generic image of the department seminar room. Visualizing the room helps me place the students in their typical locations. I can now picture enthusiastic faces, from which I infer that the reading I had assigned that week was well-received. This realization gives my simulation a layer of positively-valenced affective content, a feeling I identify with the positive momentum that would have helped me get to my hockey game that evening, rather than crashing early, as I usually do on Tuesdays. I recall the rink, my teammates, the score. I report all of this to my friend in response to her question.

Let’s stipulate that as a result of the above simulation process, I end up with an accurate mnemonic representation of what I did last Tuesday. The students and I had a great discussion of a Judy Thomson article, let’s say, and this did in fact give me the energy for late night sports. Intuitively, condition (iii) is also met: as a result of intending to remember last Tuesday, I have formed a memory that successfully refers to last Tuesday.

There is still an open question, however, of *why* (iii) is met in this case. At what point in the process—or in virtue of what—have I successfully referred to last Tuesday? Was it (a) when I initiated the recollective process via the description? Or was it (b) when the recollective process initiated by the definite description was more or less complete?

To make the choice-point sharper, between (a) referring in virtue of fixing the intended target and (b) referring as a downstream effect of fixing that intended target, we can replace the initial stipulation that my mnemonic representation of last Tuesday was accurate with a new stipulation that it was inaccurate. In more detail, suppose that the energetic class discussion, and the later events it occasioned, in fact took place the week before last. Holding fixed the same simulation process and resulting content of the mnemonic representation as in the good case, we can ask of the bad case: which occasion am I thinking of?

There are two main verdicts to adjudicate between here. One verdict is that I am referring to last Tuesday as intended but get the details wrong. We might say I “misremember”: I pick out the intended target but end up representing the events of that day incorrectly. Another verdict is that I end up with a memory of the Tuesday prior to last. We might say I’m “mistaken about what I remember”: I miss the intended target, recall (accurately) a different episode instead, but falsely believe that what I’ve simulated and described is an appropriate response to my friend’s question.

I think there is significant pressure here to issue the second verdict. I don’t merely get the details of last Tuesday wrong. I get them wrong in a very specific way, one that perfectly describes a relevant alternative episode from my past. A plausible explanation here is that the content of my intention in LAST TUESDAY was not robust enough to prevent a crucial mistake in my simulation process—i.e. visualizing enthusiastic student faces—from running interference with my initial descriptive cue and bringing in its wake a constructive process anchored to the unintended bit of content instead.

If that’s right, it reveals two things. When the gap between the intended target and the actual output of the simulation is too large, intuitions about reference track descriptive fit rather than authorial intent. Moreover, the gap is bound to be large, if the contents of the subject’s intention is to be minimal enough to ensure she doesn’t already remember what she sets out to recall. Now we’re

back to square one, as descriptive fit was the referential mechanism posited by descriptivism, a theory that, I've argued, is independently problematic.

5. Conclusion

Simulationism holds that remembering is, at root, a form of imagining. Although this idea enjoys great empirical support, it is difficult to maintain in the face of the question of how memories refer. Accounts that are fit to explain imagination-based reference in general, such as descriptivism and intentionalism, fall short as theories of memory reference specifically. Both accounts fail to be of service to simulationism because of their implications concerning knowledge of the past. The mechanism of descriptive fit treats reference as a matter of accurate representation and thus does not leave enough room for the possibility of misremembering. But referring via intending either presupposes memory knowledge or makes it look impossible to come by.

These difficulties, in turn, have broad implications for the plausibility of the constructivist paradigm more generally. In rejecting the idea that memories must preserve an appropriate connection to the past through a memory trace and replacing the causal condition with an imaginative form of contact with the past, constructivist views like simulationism set their foundations on shaky theoretical grounds. No theory of memory can be adequate that does not explain the nature of reference to, and so knowledge of, the past. The reference problem thus threatens the viability of simulationism and poses a deep challenge to constructivist accounts of memory.

Chapter Two

Generic Episodic Memories

Abstract: Remembering one's past generically is a rather ordinary feature of mental life. I can recall my daily commute to campus, from my past perspective on my bicycle, without thereby picking out one specific occasion on which I made that commute. Memories of this form have received little theoretical attention in the philosophical literature, in part because it is difficult to see how to divorce *episodic* modes of representation from representational *uniqueness*. The current paper makes progress on the question of how an episodic memory could have for its intentional object something that is not a unique episode. It proposes an analysis of generic episodic memories on which the remembering subject represents an event *type*. Genericity over episodes is best explained in terms of the subject's knowledge of what it was like to experience events of that type.

1. Introduction

In classifying acts of conscious recollection, here are two questions we might ask. First, we might want to know what it is that one remembers. What kind of thing (event, object, property, fact) does a subject's memory relate her to? Call this the *intentional object question*. Second, we might want to know something about the subject's perspective while remembering. In what representational mode are the objects of her memory presented? Call this the *representational format question*.³⁸

³⁸ See Siegel and Silins 2017 for a similar way of carving out the theoretical space. They distinguish questions about the *memory relation* (e.g., when I remember the past, what relatum stands on the other end of mnemonic relation from me?) from questions about the *memory perspective* (e.g., what is phenomenally distinctive about this form of consciously recollecting?)

It is now fairly standard to distinguish two kinds of memory, *episodic* and *semantic*, in terms of answers to these two questions. On the intentional object front, episodic memories are commonly thought to be about previously experienced events from one's past. Semantic memories, by contrast, can represent general information or conceptual knowledge, which—although such information may be acquired from multiple events—need not be tied to any localized spatiotemporal context.³⁹

Episodic and semantic memories are also typically differentiated by their representational formats. Episodic memories represent their objects by way of phenomenally rich, imagistic contents, and typically feature autonoetic phenomenology, that is, the feeling constituted by an awareness of oneself in the past. Semantic memories instead represent their objects discursively (for example, via propositional contents) and exhibit no distinctive recall phenomenology.⁴⁰

The two dimensions of difference can be illustrated with the following example. Compare how I might remember the festive occasion of my sixteenth birthday party, via perceptual imagery of chocolate cake, my parents' backyard, and certain familiar faces, with how I might remember that birthday parties are generally festive gatherings, without thinking back to any occasion in particular or conjuring up any imagery.⁴¹ The former act of recollection is episodic; the latter is semantic. Or so the standard taxonomy has it.

³⁹ This way of drawing the distinction is originally due to Tulving 1972. To clarify, the semantic memory system is also responsible for the retention of specific bits of information, such as my knowledge that my parents got married on the coldest day in recorded Saskatoon history.

⁴⁰ There are myriad characterizations of the paradigmatic phenomenology of episodic memory (including the feeling of 'replaying' events, 'mental time travel', familiarity, or pastness), as well as disagreement over whether the phenomenology is part of the memory state itself or the result of a metacognitive monitoring system (see Teroni 2017 for a helpful discussion). The present point is that episodic memories paradigmatically feature *some* kind of feeling that the remembered object is located in one's past, in contrast to the recollection of facts (though see Frise 2018 for a discussion of 'tip-of-the-tongue' cognitive phenomenology in connection with the *forgetting* of facts).

⁴¹ This characterization is meant to be ecumenical, as far as possible, on a number of core debates: whether the contents of episodic memory are preserved vs. reconstructed, whether they are alethic vs. authentic, and whether perceptual imagery involves low-level or high-level contents.

Although theoretically tidy, the standard way of carving up the declarative memory territory leaves little room for a familiar form of remembering: memories of general, or repeated, events. Here is an example of the phenomenon I have in mind:

The Commute. Suppose you ask me what my daily bike commute to the office was like when I lived in Boston. I think back to that period of my life, and a familiar scene unfolds in my mind's eye: the gleam of the river in the morning light to my left, the feeling of pedalling hard up a steep incline, being on high alert for unpredictable drivers, and taking the final turn into the university grounds as I dismount. I recall this bike ride with ease, familiarity, and faint nostalgia.

The memory described in *The Commute* exhibits a puzzling combination of features. On the one hand, my memory seems to have an episodic mode of presentation. In recalling my regular bike ride to campus, I am recruiting phenomenally rich contents, elements of which are visual (the look of the river), kinaesthetic (the sensation of peddling), and affective (the anxiety of being a cyclist amongst motorists). I am representing the unfolding of a scene as though from a point of view I once occupied—that is, my memory contents include egocentric spatial relations, or what some call spatial-perspectival imagery (the river is presented to my left). And there is a sense in which, as the scene unfolds, I am mentally re-playing or imaginatively simulating something I previously experienced. We can call this combination of features *episodicity*.

On the other hand, my memory seems to share with the category of semantic memories in that it represents information derived and synthesized from multiple occasions, and in not being indexed to a particular, temporally local event. Since I took the same bike route almost every weekday, I had multiple past experiences with very similar contents. When it comes to the causal connection

between my mnemonic representation, and my past activities of commuting, then, it is likely that multiple causal chains are explanatorily relevant. And if, in response to my memory report as described in *The Commute*, you were to ask me which particular ride to campus I am recalling, I would be at a loss as to know how to answer you: I might reply, “I don’t know,” “none in particular,” or “all of them, in a sense.” This is because, I suggest, the contents of my memory form a general representation of my past commutes. We can call this feature of my memory *generic representation*.

I will call memories that exhibit these two features *generic episodic memories*. The phenomenon has been relatively undertheorized in its own right, mostly appearing in passing connection with discussions of episodic memory as a related phenomenon or a borderline case.⁴² But examples are easy to come by. You might recall learning to swim over the course of several summers, meetings with your advisor in graduate school, or migraines you used to get as a child. In each of these cases, the memory captures something that is unified across a range of individual past occurrences, and yet presents the past with some level of experiential detail. As a first-pass description of the account I will defend, generic episodic memories are a form of experiential remembering that represent event *types* rather than particular events. As such, they refer generally rather than uniquely to one’s personal past, despite having an episodic mode of presentation.⁴³

⁴² Byrne 2010, for example, presents a case of recalling what skunks look like, whereby one calls to mind a visual image of a prototypical skunk that is based on various skunks one has seen before, as an apparent counterexample to the standard claim that episodic memories are about personally experienced events. Hopkins 2014 goes further, providing a definition of generic memories: “memory of how things present themselves in experience, but not how any particular thing presented itself in any particular experience” (p. 326). Hopkins claims that the absence of a particular representational object renders generic memories non-episodic, but also insists that they exhibit the same “range of expression” as episodic memories—e.g., the representation of the past to oneself via a mental memory image. Hoerl 2014 provides a theoretical account of the distinction between remembering looks and remembering (particular) events. I postpone discussion of Hoerl’s approach to section four, where I argue that memory of what something looks like does not fully capture the phenomenon I have in mind (though it comes closest). Other discussions of general event memories include Andonovski 2022, Campbell 2002, Conway 2005, and Openshaw 2022.

⁴³ In cognitive science, related theoretical notions include “schemata” and “scripts.” Schemata are theorized as cognitive structures responsible for organizing information about items or events, which have been invoked to explain, for example, how episodic memories become less detailed and more gist-like over time. (See Brainer & Reyna 2002; Mori 2009; Robin and Moscovitch 2017; see also Schurgin 2018 for a helpful review.) Scripts are structurally similar to schemata in organizing information, but specifically concern the sequences of behaviors or expectations that are relevant to action-guidance in

The term ‘generic episodic memory’ might strike the reader as an oxymoron: episodic memories are, by definition, about episodes, and since episodes are temporal particulars, there must be a fact of the matter as to which episode any given memory represents. Our first order of business will thus be to carve out the conceptual space for the phenomenon within the current taxonomy of memory. I do this in §2, in which I also explain where the philosophical pressure to pair episodicity with representational uniqueness comes from, as part of a diagnosis for why the generic variety of episodic memories has been underexplored.

The remainder of the paper is structured as follows. In §3, I motivate the core theoretical difficulty in countenancing memories that are both generic and episodic, by presenting a puzzle about how generic episodic memories could be both accurate and authentic. In §4, I offer my positive account, the event-type theory of generic episodic memories, and show how it satisfies the explanatory task uncovered in §3. In §5, I explore two alternative approaches to my account, each of which attempts a reductive analysis of one of the core features of generic episodic memories. I argue that neither approach is theoretically satisfactory and hence that the phenomenon is not assimilable to other forms of remembering.

2. Taxonomical resistance

In both the philosophical and psychological literature, episodic memories are typically characterized as representing specific events or experiences. The following is a small but representative sample:

certain kinds of situations. For example, one might have a script for dining at a restaurant (Abelson 1981). Neither theoretical notion, however, is identical to generic memories. In the case of schemata, these are discussed in the context of explaining the content of singular episodic memories. In the case of scripts, these are discussed in the context of action and procedural memory. Our topic, by contrast, is the episodic representation of one’s general past.

The fundamental feature of episodic memory: the fact that episodic memories are memories about specific past events. (McCormack 2001, p. 287)

Episodic memory is, by definition, unique to a single event. (Squire et al. 2004, p. 286)

Episodic memory... is necessarily about autobiographical events that have a singular spatiotemporal location, whereas semantic memory may very well be about general facts (Perrin 2019, p. 481, translation mine)

This operating idea is arguably traceable to Tulving's original distinction between the episodic and semantic memory systems, which turned on whether the retrieved information is linked to a specific spatiotemporal context of acquisition (Tulving 1972). But the link between episodicity and specificity is not merely taxonomical; it has also become entrenched for philosophical reasons.

One reason to think that episodic memories represent uniquely and not generically comes from a general theoretical approach to memories as mental representations. At the heart of this approach is the idea that various modes of remembering are closely linked to the nature of the remembered objects. The basic idea is to account for a distinctive feature of the representational vehicle in terms of some property of the represented object, in the way that, for example, the spatial properties of a map are explained in part by the spatial properties of what the map depicts.⁴⁴ By authorizing an inference from representational mode to represented object, explanations of this form can seem to suggest that episodic memories represent unique episodes. Here are three examples from the literature.

⁴⁴ I am borrowing this example from Teroni 2017. A similar point holds of iconic representational formats in general.

Replay phenomenology and temporal structure

Suppose we begin with an observation about the representational format of episodic memories. As noted in section one, one respect in which my memory is episodic is that it exhibits what is often called “replay phenomenology.” When I recollect, it feels as if I am re-living a scenario from my past, supplying my experience of remembering with a similar temporal structure to that of my original experience. In *The Commute*, for example, I imagine myself making my way to campus, the route unfolding in accordance with my first pedalling away from home, then in a certain direction along the river, and finally arriving at the office. This feature makes my representation of the past less like an act of judgment (that such-and-such occurred) and more like a mental episode that itself unfolds over time, in much the same way that episodes themselves do.⁴⁵

So, if pressed to explain why episodic forms of remembering have replay phenomenology and thus temporal structure, whereas other forms of remembering (like memory of facts) do not, it would be natural to appeal to a difference in their respective objects. Particular events and experiences, but not facts or general knowledge, have temporal structure. They are the kinds of things that can be “replayed,” because they are the kinds of things that happen over time. Thus, events or experiences are what episodic memories are about.

Cognitive contact with events

Martin 2001 employs a similar strategy to distinguish episodic memory from the recall of facts. But rather than working from a distinction between modes of presentation as from the subject’s point of

⁴⁵ An idea like this can be found in Soteriou 2013. In distinguishing episodic from factual forms of recollection, Soteriou appeals to the structurally similar temporal profiles of episodic memories and perceptual experiences: “An act of episodic recollection appears to have a temporal profile that is similar to one’s original experience of the event recollected. One’s episodic recollection of a past event unfolds over time in a way that is similar to the experience one had of that event, and this isn’t true of an act of the semantic recollection of some fact about that event” (p. 171).

view, he appeals instead to a difference in two modes of cognitive contact between subjects and objects: apprehension and knowledge. Apprehension is episodic in nature: “seeing, feeling, tasting something are all events or occurrences” (p. 265). Knowledge, by contrast, is a standing condition or static state. Because events are temporal particulars, the nature of our cognitive contact with events must be episodic. Events are therefore not known, but are apprehended, such as through perception (266). Due to this restriction on mode of contact, Martin argues that “episodic memory, then, traces back not to past knowledge, but past apprehension” (p. 265), from which he infers the proper objects of episodic memory are particular events. Factual memory, by contrast, consists in the preservation of knowledge, whose objects are facts.

Representational dependency

Perhaps the most promising version of the present strategy is an argument for why episodic memories are memories of past experiences, as opposed to unmediated memories of events themselves. Roughly, the argument is that positing past experiences as the primary representational objects of memory best accounts for the following fact about episodic recollection: its phenomenal character is akin to, but not quite the same as, the phenomenal character of perception.⁴⁶

One way to account for this contrast—that is, to explain both the similarities and the differences in phenomenal character between the two kinds of conscious experience—is in terms of a one-way representational dependence. Specifically, memory includes, as part of its total contents, past perceptual experience.⁴⁷ When I visually recall a purple flower, for example, I am recalling a past

⁴⁶ Candidates for the relevant phenomenal dissimilarities include temporal properties, degree of assertoric force, vivacity, content coherence, and contextual detail.

⁴⁷ Defendants of such a claim include Searle 1983, Martin 2001, and Soteriou 2013, though their accounts vary substantially in their details.

visual experience of that flower. My memory contains my past perception as a representational object, and thereby inherits the phenomenal character and contents of my past perception. Thus a feature of the mode of representation—my memory’s phenomenal character, as contrasted with that of perception—is explained by positing a proprietary representational object—my past perceptual experience.⁴⁸

In each of the three examples above, an aspect of the memory qua mental representation—(i) its temporal structure, (ii) the kind of cognitive contact it affords, and (iii) its phenomenal character—is explained by some feature of the represented object—(i) the temporal structure of events, (ii) that events but not facts are apprehended, and (iii) the comparative phenomenal character of perceptual experience. This form of explanation permits an abductive inference from intentional mode to intentional object. In all three cases, the inference is from episodic remembering to a remembered object that is *singular*, such as a past event or experience. The availability of this kind of inference is one reason why representational uniqueness may be thought to hold for episodic memories as a class.⁴⁹

Although the connection between representational modes and represented objects is easy enough to motivate, it does not hold as a general rule for mental representations. An obvious difficulty is extension to non-experiential representations. Sentences, for example, have discursive rather than

⁴⁸ The idea that one phenomenally conscious mental state takes another for its representational object was first raised for experiential forms of imagination (cf. Vendler 1984 and Peacocke 1985). Martin 2002 later introduced the Dependency Thesis, which he formulates as follows: “to imagine sensorily a Φ is to imagine experiencing a Φ ” (2002: 403). For an extensive discussion of how the principle of representational dependency applies to both perceptual imagination and episodic recollection, see Soteriou 2013.

⁴⁹ There may, of course, be other reasons why uniquely referring episodic memories have been the focus of theoretical attention. As an anonymous referee helpfully points out, uniquely referring episodic memories are perhaps central to personal identity. For instance, such memories pick out singularly important episodes, or defining moments, from one’s personal past. Yet generic memories may also be central to personal identity, albeit in a different way. Plausibly, by depicting types of events that were typical of a chapter in one’s life, they play an important role in the narrative structure in which the more defining moments are embedded (cf. Conway 2005).

iconic representational structure, which is not straightforwardly explainable in terms of the properties of the objects represented by a sentence.

Moreover, the “episodic mode, episodic object” rule fails in the other direction as well. Particular events, which are episodic, can be represented through non-episodic means, such as beliefs or other propositional attitudes. For example, I might regret having gone to a party last night, where my state of regret is a standing attitude of mine towards some past happening—my going to the party—without itself being an episodic mode of representation.

The mode-to-object inference does not even generalize to other mental representations with phenomenally rich, experiential contents. Take the paradigm case: perception itself. Perceptual experiences are episodic. Nevertheless, much of what perceptual experiences represent—e.g., low-level properties like colours and shapes, high-level properties like kinds or affordances, and the people, objects, or scenes these properties belong to—are not themselves episodic, but rather persist across events.

Clearly, the underlying inference from mode to object is not valid. It doesn’t follow from the fact that episodic memories are experience-like qua mode of representation, plus the fact that experiences are particular occurrences, that episodic memories thereby only represent particular occurrences. For instance, a different way to accommodate the relevant phenomenological datum driving the inference—i.e., that episodic recollections are in some important way like perceptual experiences—would be to countenance whatever the objects of perception are as possible objects of memory.⁵⁰

⁵⁰ Openshaw 2022, for example, defends a form of experiential memory for objects themselves. I discuss his view in §4. But defenders of representational dependency might not object to including the objects of perceptual experience in the total representational contents of memory. Supposing perception is transparent, then, plausibly, memory will inherit that transparency. This kind of move would make sense of the way in which the contents of one’s memory of seeing the northern lights, for example, seems to include the northern lights themselves as a represented object.

Nevertheless, proponents of dependency will still be pushed towards positing two tiers of representational objects: memories represent perceptual experiences primarily, and only by representing the experience do they also represent whatever it was that experience represented. The distinction in tiers will then account for why the phenomenal

De-pairing episodicity from episodes thus creates theoretical space for the idea that there might be episodic memories that represent the past generically. However, this still leaves much to be explained, including how to make positive sense of the combination of episodicity and genericity. In the next section, I will argue that, taxonomical resistance aside, this task isn't easy.

3. The puzzle of accuracy and authenticity

Suppose we begin with an epistemological question about episodic memory: What does it take to remember the past well, i.e., well enough to count as knowing about what happened by remembering it?⁵¹ An intuitive response is that one's memory must faithfully represent one's past. Following Bernecker 2010, we can separate two respects in which one's memory can serve this representational function.

First: my memory must represent the way things actually stood in the past. When it does so, my memory is *accurate*. For example, if I recall a red sedan blowing through a yield sign at a certain intersection, and a red sedan did in fact fail to heed a yield sign at that location, then my memory accurately depicts the relevant incident. Second: my memory must represent the way I previously experienced things having stood in the past. When it does so, my memory is *authentic*. For example, if I recall the red sedan blowing through a yield sign from the visual perspective of my office window

character of memories and imaginings are different from the phenomenal character of perception, which was part of the original motivation for the view.

⁵¹ I intend for the locution 'knowing by remembering' to be read ambiguously between remembering as a *basis* or *source* of knowledge and remembering as a *kind* or *way* of knowing, so as to remain neutral between non-factive accounts of memory (cf. Michaelian 2016) and factive accounts of memory (cf. Williamson 2000).

overlooking that intersection, and I did originally witness the scene from that spatial location, then my memory authentically depicts the relevant incident.

According to Bernecker, both accuracy and authenticity are necessary for knowing by remembering: “a memory state must accord not only with objective reality but also with one’s initial perception of reality” (2010, p. 214). When one’s memory state meets the former condition but fails the latter, it is in some important sense distorted.⁵² Faithfulness to how one previously experienced the past is as relevant to the epistemic success of one’s memory as is faithfulness to the past per se.

The notion of authenticity is meant to apply exclusively to episodic memory.⁵³ This is because episodic memory has a particular connection with past perceptual experience—or so many theorists suppose. The idea has its roots in the preservationist tradition of memory. Preservationism dates back at least as far as Aristotle (*De Memoria*), who thought of remembering as perceiving an image causally derived from the past. Since then, various roles for memory in connection with past perception have been proposed. Memory is thought to retain past perception (Martin 2001) or the cognitive contact supplied by past perception (Byrne 2010), to represent past perception (Searle 1983), and to replicate past perception by retaining its contents (Martin and Deutscher 1966; Recanati 2007) or by reconstructing them (De Brigard 2014; Michaelian 2016). Even theorists working squarely within the constructivist tradition tend to think of episodic memories (viz. imaginative simulations) as embodying

⁵² The paradigmatic example of an inauthentic memory is an *observer memory*: an episodic representation of a previously experienced event from a perspective different from the one the subject actually occupied at the time of the remembered event. Recently, however, many theorists have defended the view that observer memories should count as authentic after all. See McCarroll 2018 for a sophisticated version of such a defense.

⁵³ Extending the notion of authenticity to semantic memory would be difficult. Consider an analogous proposal: to count as knowing a proposition by remembering it, the proposition must both be true and previously believed by me in the same way (i.e., based on the same evidence) that I now believe it. But what if I now believe the relevant proposition on the basis of memory alone rather than the original source of information? A large quantity of one’s beliefs (including, e.g., the class of beliefs about trivia) exhibit some such basing structure and would thereby not qualify as knowledge. Authenticity for semantic memory thus seems much too stringent an epistemic condition.

experiential perspectives and representing experiential content that *could have* featured in (though need not be causally derived from) a particular past experience.

Differences aside, a shared intuition of these theorists seems to be that one's past perceptual representation of the world is what one's later recollection aims to reproduce for the subject (through either a preservative mechanism, like a sustaining memory trace, or a constructive mechanism, like mental simulation). The notion of authenticity captures this intuition by emphasizing accordance with one's previous experiential perspective onto the world, as distinct from accordance with objective reality. It thus also accounts for the *episodic* features of memory's representational format: phenomenally rich contents, a point of view, and replay phenomenology.

Applying the joint notions of accuracy and authenticity to *generic* episodic memories, however, is much more theoretically difficult. Can my memory, as described in *The Commute*, represent my past both accurately and authentically—i.e., both as things really happened and in the way I previously experienced things happening? As I will now explain, although generic episodic memories can plausibly be seen as either accurate or authentic, it is difficult to hold both features simultaneously in view. This will reveal why the combination of *generic representation* and *episodicity* is, indeed, puzzling.

Consider the accuracy question. What would it take for my memory in *The Commute* to be accurate? The answer to this question seems to depend on considerations such as the following: whether in fact I did regularly bike to campus, taking more or less the same route, and often during rush hour. If so, then the contents of my memory get things generally right. Assessing my memory for its accuracy, in other words, requires treating my memory as a *general* representation, not measured against this or that occasion of commuting but by the general types of features common across them.

However, considered as a general representation of my past, it is hard to see how my memory could possibly be *authentic*. My memory, it seems, represents my past in a general way, via type-level features, yet *I* did not previously experience these past events in a general way. I experienced concrete,

event tokens, such as individual bike rides, that take place on temporally specific occasions. Seen in light of its generic aspect, my memory in *The Commute* cannot represent things in the way I previously experienced them.⁵⁴ It may be accurate, but it won't be authentic.

We can also observe a similar tension running in the opposite direction. Suppose we first consider the authenticity requirement instead. Does my memory in *The Commute* present my past in the same *way* I previously experienced it? There is strong reason to think that it does. When I inhabit my generic memory, I represent my past as though I were experiencing it again. My memory has, after all, an *episodic* character. It features the sensation of pedalling, the felt presence of traffic, and the look of my surroundings from my spatial perspective on the bicycle. Moreover, when I inhabit my generic memory, these features might get *filled out* in a particular way: for instance, the chain of my bicycle feels smooth, the traffic around me contains a lot of trucks, and the river is reflecting the sunshine. These experiential details are part of what makes my memory authentic—they are part of what makes it the case that my memory presents my past with the same particularity with which I experienced it.

But suppose it was just as often cloudy as sunny on these rides, or that my bike occasionally had mechanical issues, or sometimes there were no trucks on the road. In fact, there may have been no single past event—let alone many such events—that corresponded to the particular combination of features that my memory presents me with. If we judge the accuracy of my memory on the basis of the particular details that show up when I mentally 'replay' the experience of commuting, it seems we are forced to conclude that my memory does not represent my past very well. It may be authentic, but it won't be accurate.

The dual aspect of generic episodic memories thus raises a puzzle. On the one hand, when we focus on their generic aspect, we are able to make sense of their being accurate, but not of their being

⁵⁴ A similar argument is made in Openshaw 2022 against what he calls 'Experientialism,' a view on which to experientially recall x is to recall an experience of x . He argues that the experientialist thesis cannot accommodate certain substitutions for ' x ', including general events (p. 15).

authentic. On the other hand, when we focus on their episodic aspect, we are able to make sense of their authentic, but no longer of their being accurate. The core theoretical difficulty is how to conceive of generic episodic memories in a way that would enable us to make sense of their being *both* accurate and authentic at the same time. Solving this puzzle is important if we are to understand how generic episodic memories could provide us with knowledge of our past.

At the most general level, the explanatory burden on any account of generic episodic memories seems to be the following: account for the felt particularity of the memory representation, plus the generality with which that representation nevertheless ranges over one's past experiences. In the next section, I demonstrate how my analysis of generic episodic memories meets this explanatory demand.

4. The event type theory of generic episodic memories

This paper began with two questions about memory: the intentional object question (what kind of thing does one's memory represent?) and the representational format question (in what mode does one's memory represent that thing?). I propose we analyze generic memories as a kind of experiential memory that represent event *types* rather than event instances from one's past. In *The Commute*, the intentional object of my memory is a kind of event I repeatedly participated in, namely, riding my bike to campus on an ordinary weekday. That event type is represented in an experiential mode, via contents abstracted from multiple past experiences of commuting. These contents express what it was like for me to make that commute.

My account captures generic representation. Because my memory represents an event-type—commuting to campus—it generalizes over the relevant set of past events without singling out one in

particular. This is what makes a request for referential specificity inappropriate, for what I am remembering is not biking to campus on this or that occasion but doing so in general.⁵⁵

My account also captures episodicity. In *The Commute*, I recall what it was like to bike to campus by means of phenomenally-rich imagery, such as the visual appearance of the river, and the sensation of pedalling. The contents of my recollective experience are presented as though from the point of view I used to occupy on my bike, one which I reconstruct from the spatial-perspectival information abstracted from my past experiences. When I recall commuting from this perspective, I mentally rehearse a paradigmatic example of the event type, which gives my memory its replay feel.

The analysis I am proposing also helps us see how memories with both of these features could nevertheless be both accurate and authentic. That is, it gives us the theoretical tools to explain how these memories could be at once (1) episodic yet accurate, as well as (2) generic yet authentic.

The episodic mode of presentation is what can seem to give generic memories a felt particularity. Because the perspective I occupy in memory is a representation of the perspective I occupied during any particular past episode, the experiential contents I recruit to re-create the remembered scene as from that perspective are potentially detailed or granular enough to suggest I have a particular past experience in mind. The natural thought here is that the more I fill things in, when recounting my commute, the more I am thereby excluding from my memory's representational umbrella experiences of the commute that don't quite fit the mold.

⁵⁵ One emerging view of the causal architecture of memory storage, the distributed trace model, lends empirical backing to the idea that memory contents can generalize over multiple episodes from a subject's past. On the distributed trace model, the information acquired from any particular episode is spread across a neural network rather than stored in a separate file. As a subject's experiences accumulate over time, similarities in contents give rise to activation patterns between semantic nodes in the network. On this model, memory traces are best thought of as the preservation of experienced patterns across events rather than the retention of discrete representations of particular events.

The first account of distributed trace theory was developed by Sutton 1998; see Robins 2016 for critical discussion of different alternatives. Michaelian (2011) and De Brigard (2014) both adopt the basic framework for their respective versions of simulation theory.

But, in fact, it doesn't follow from granularity of contents that my generic memory thereby represents uniquely. Prototype representations are a case in point. Take, for example, a drawing of an oak leaf in a textbook that a student uses to learn to identify leaf types. The drawing depicts a *type* of leaf, and thus represents oak leaves in general, despite perhaps having been once traced from and thus inheriting the concrete configuration of one such leaf. Similarly, I might remember a type of event by means of a specific prototype, such as when I recall in great detail watching my favorite film (which I have seen countless times) without thereby individuating a particular occasion on which I watched it.⁵⁶ It is thus a mistake to infer representational uniqueness from the seeming particularity of contents.

This observation helps us resist the horn of the dilemma in §3 on which episodicity, in virtue of inviting some element of specificity, precludes accuracy. The worry, recall, was that if I represent my bike rides through traffic by, say, picturing trucks on the road specifically, this will render my memory inaccurate with respect to particular bike rides on which I did not encounter trucks. We are now in a position to see how that kind of accuracy assessment is a mistake: it measures a general representation against particular instances in the way we measure *particular* representations against particular instances. Compare a uniquely referring memory: if I am remembering my commute *yesterday*, and if on that day I only encountered cars, then a representation of my commute in which I encountered many trucks would not be accurate. But in the case of my generic memory, picturing trucks is just a specific way of representing traffic that was salient across many experiences of commuting. The fact that my generic memory foregrounds trucks as part of its content does not seem to preclude its getting things right in general.

⁵⁶ See Langland-Hassan's (2022) discussion of the phenomenon of 'forgetting through repetition' for an argument for why one's recollection of the film would not pick out any particular event of watching it. His own verdict about this case is that, since one does not have a uniquely referring episodic memory of watching the film, one does not *remember* the film at all, but merely accurately represents it through experiential imagination. In drawing this conclusion, Langland-Hassan has overlooked an important alternative, viz. the subject has a generic episodic memory of the film.

Of course, it remains an open question how exactly phenomenally rich representations of event types are to be assessed for accuracy, given that there will tend to be variation in the set of token events belonging to the type. Take, for instance, the feeling of wariness of motorists with which I recall my commutes. In order for my memory to be accurate, does it need to be the case that this feeling was present (or present to the same degree) on each of my rides? Plausibly, exceptions should be permitted, in much the same way that generic claims about members of a kind—such as “Ravens are black”—are true despite some or even many instances of the kind not possessing the relevant feature.⁵⁷ For instance, perhaps my memory is accurate because it is sufficiently likely that, for any randomly selected instance of commuting, I felt that anxiety. Or perhaps my memory is accurate because the times at which I felt relaxed in spite of the traffic are exceptions to the “norm”.⁵⁸ Determining which, if any, of these models provides the correct accuracy conditions for generic memories is beyond the scope of the present paper. But what seems clear is that the comparison to generic statements will provide an instructive starting point.

Let us now turn to the other strand of the puzzle, concerning how a memory could be generic yet authentic. The second difficulty was understanding how, in recalling something general, a subject could be recalling her past in the same way she experienced it, namely, as the participant of particular events. On my event-type theory of generic memories, this difficulty begins to dissolve. A key part of the proposal is that a remembering subject represents event types via contents that encapsulate *what it was like* for her to experience events of that type. This notion—recalling what something was like—is the key theoretical tool to unite genericity and authenticity. This is due to its dual aspect. On the one hand, it is a distinctively phenomenal or experiential concept: one knows that something is like insofar

⁵⁷ In this way, generics are distinct from universal generalizations, such as “All ravens are black.”

⁵⁸ These two suggestions for how to model the accuracy conditions of generic memories are inspired by statistical and normality views of generics, respectively. See Leslie (2012) for helpful overview of the generics literature.

as one has experienced it first-personally. On the other hand, the concept naturally invites type-level readings. We often speak as though there is something it is like, in general, to fall in love, go on vacation, or to feel nervous in front of a crowd.

But what exactly does remembering what something is like consist in? My account of this notion draws on extant literature on *knowledge* of what something is like. There are two main approaches in the literature: a dispositional approach, and an informational approach. As I will explain in the remainder of this section, my account of generic memories can be further developed within either of these two theoretical frameworks.

The starting point for the dispositional approach to remembering what something is like, is the connection between remembering *how* to do something, such as ride a bicycle, and remembering *doing* that thing from a first-personal perspective. For example, both kinds of memory might activate the same sets of sensory-motor dispositions. Similarly, many analyses of *knowledge how* (the relevant contrast class being *knowledge that*) link a subject's possession of know-how to their ability to represent certain states of affairs under a practical mode of presentation, such as experiential memory or experiential imagination.

Lewis (1990), for example, argues that knowing what an experience is like consists in possessing a suite of cognitive abilities that includes being able to imagine having that experience, episodically remember having that experience, and identify future tokens of that experience. Importantly, these abilities are not grounded in propositional knowledge but a complex set of sensory dispositions or recreative capacities. For example, one knows what red is like if one is disposed to recognize the color when one sees it again, or if one can imaginatively reconstruct or recall a red object.

Extending Lewis's ability analysis to *The Commute*, what we could say is that my experiential knowledge of what it's like to commute to campus consists in my ability to imaginatively recreate, or recall, an experience of that event-type with the same phenomenal character. We can analyse the

components of the phenomenal character further to include experiential knowledge of riding a bike, dodging traffic, looking at the river, and so on.

A potential epistemological worry with the present analysis of knowing what it's like is that it will turn out to be circular when applied to memory. For, earlier it was claimed that my memory activates or expresses my knowledge of what it's like to experience a type of event, and now that very knowledge is being analyzed in terms of the ability to remember that event type.

One way to address this concern is to lean on the distinction between standing and occurrent states of knowing. Soteriou's (2013) theory of episodic memory, for example, trades on the connections between the subject's standing knowledge, occurrent knowledge, and a phenomenally conscious episode. On his view, a subject's standing state of knowing what it was like to "apprehend" (experience) some event consists in the ability to put oneself into an occurrent state of knowing what it was like to apprehend that event, where the occurrent state of knowing is constituted by the phenomenally conscious episode of recollection. Where Soteriou's view departs from the pure ability analyses of Lewis's, however, is in treating the contents of one's occurrent knowledge *propositionally*. Soteriou says:

One knows what it was like to apprehend some past event in the sense that one has an ability to do something that can put oneself in an occurrent state of knowledge whose propositional content is a distinctive kind of answer to the question, 'what was it like to apprehend that event?'. (2013, p. 176)

To connect the possession of an ability to the activation of propositional knowledge, Soteriou deploys the framework developed by Williamson and Stanley (2001) according to which S knows how to Φ just in case (i) S knows *that* W is a way to Φ , and (ii) S entertains this proposition under a practical

mode of presentation (p. 430). In the case of memory, the phenomenal character of the episodic recollection meets both conditions, because it supplies the contents of W (a way of experiencing an event) and represents those contents as from a first-person point of view characteristic of the episodic memory perspective.

Soteriou therefore avoids the possible circularity worry facing the pure ability analysis of the knowledge of what it's like that is activated in memory by positing that the subject retains some propositional information, viz., information of the form $\langle W \text{ is a way to } \Phi \rangle$. So, even though the remembering subject's standing knowledge is connected to a suite of capacities or dispositions to form the right conscious imagery, it is still grounded in the retention of factual knowledge.

Another approach to knowledge of what it's like that grounds the ability to experientially recall in the retention of information is due to Papineau 2006. Following his view, we might analyse the remembering subject's standing knowledge of what F is like in terms of the subject's possession of the perceptual concepts that enable her to recognize or recreate an F-experience. The information associated with each perceptual concept is stored in what he calls a "sensory template" (p. 115). These templates are first created through perceptual experiences and are updated through subsequent activations (consisting in a sensory match, or "resonation" between the stored pattern and the exogenous or endogenous stimuli). For example, if I know what a robin's birdsong sounds like, that is because I have stored that sensory pattern in the file indexed to my perceptual concept for $\langle \text{robin} \rangle$, and if my memory activates that knowledge, that is in virtue of a match between the phenomenal character of my recollective experience and my underlying sensory template (p. 118).

Both Soteriou's and Papineau's analysis of the knowledge of what it's like that gets activated in an episodic memory can accommodate genericity. Papineau's perceptual concepts come in "structural hierarchies" which enable representations of, for example, both bird-tokens and bird-types (p. 117). Applying his framework to my view, we could analyze the contents of my generic memory

in *The Commute* in terms of perceptual concepts (for experiences of biking, anticipating traffic, looking at the river) whose sensory templates get co-activated when I recall the event-type of that commute.

On Soteriou's analysis, the structure of the proposition that the remembering subject is related to, <W is a way to experience E>, lends itself to different possibilities for genericity. For one, the character of W could be individuated in terms of phenomenal types. Alternatively, W could relate to E in the way a prototype relates to the type exemplified, as captured by the relation 'is a way to experience'. The remembering subject's grasp of the total proposition <W is a way to experience E> would then structure the memory perspective such that the subject recognizes that its contents represent the past generically (i.e. depict an event-type) rather than uniquely (i.e. depict an event-token).

The event type theory of generic episodic memories is therefore amenable to at least two varieties of analyses of knowing what something is like.⁵⁹ Both of these approaches *informational approaches* in that they ultimately ground knowledge of what it's like in the grasp of propositions or the possession of concepts.⁶⁰ Because natural interpretations of the latter two entail that the subject has stored the relevant information from the time of her past experience(s) to the time of her present recollection, such analyses will not be particularly appealing to a wide variety of constructivist theories of episodic memory. Before closing the present section, I want to gesture at a way my account could accommodate a dispositional approach as well, such as Lewis's ability analysis of experiential knowledge.

⁵⁹ Both views also have the conceptual resources for explaining cases of *misremembering*, i.e., generic episodic memories with nonveridical contents. Soteriou could hold that the misremembering subject is related to a false proposition, one that does not constitute an answer to the question of what it was like to apprehend a kind of past event. (In more detail, the subject activates a belief in a proposition of the form <W is a way to experience E> but where the phenomenal character with which W is individuated does not pick out a way of experiencing event type E.) And Papineau could analyze cases of misremembering in terms of an activation error, such as a mismatch between the perceptual concepts that are co-activated during recollection and the perceptual concepts that would in fact be required for a representation of the target event type.

⁶⁰ See McGrath 2017 for an illuminating discussion of "intellectualist" versus "anti-intellectualist" approaches to experiential knowledge, a division which maps onto what I am calling "informational" versus "dispositional" approaches.

The circularity worry raised earlier for the pure ability framework as applied to memory went as follows: the memory theorist cannot both (i) analyse episodic recollection in terms of the activation of experiential knowledge, and (ii) analyse experiential knowledge in terms of the capacity to episodically recall. It is worth noting, however, that the concern only has purchase when the notion of knowledge activation is given a preservationist reading—that is, if we suppose the remembering subject’s knowledge of what it’s like is something that they have retained all along. But on certain theories of memory, the knowledge one’s episodic memory provides is something that is *generated* when one accurately represents an episode from their personal past.⁶¹ On the generative reading, there is no equivalent worry that a dispositional analysis of the subject’s *standing* possession of experiential knowledge will turn out to be circular, since the relevant knowledge is constituted only through the occurrent exercise of the subject’s capacity to remember.

Combining the generative understanding of memory knowledge with a dispositional approach to experiential knowledge, we get the following analysis: generic episodic memories activate experiential knowledge when the remembering subject exercises her disposition to episodically construct or simulate an event-type. That disposition, furthermore, is grounded in the possession of a properly functioning episodic construction system and access to a distributed network of memory traces.⁶²

⁶¹ Michaelian 2011 first proposed epistemic generativism for memory constructivism. An earlier proponent of the idea that memory is a generative epistemic source is Lackey 2005. Bernecker 2010 defends a version of generativism for a causal theory of memory, and Fernandez 2016 defends a version for a self-reflexive account of memory contents. But see also McGrath 2007 for a discussion of the ‘epistemic boost problem’ and a defense of epistemic conservatism.

⁶² Simulationist theories can also accommodate type-level contents fairly easily, as a result of their commitment to the continuity between episodic remembering and other experiential forms of imagining, such as imagining a counterfactual or hypothetical future episode. Imagining future events has been analysed (cf. Debus 2014) as the sensory imagining of an event type embedded within the subject’s temporal beliefs (such as the belief that a token of this type is located in the future). A similar analysis could be given for generic episodic memories as imaginatively constructed event types projected onto the past. *Nonveridical* generic episodic memories, moreover, would be due to an error in the process of episodic construction (which, on simulationist views, are par for the course in a healthy functioning memory system; see De Brigard 2014), resulting in an inaccurate simulation of what it was like to experience an event of that type.

What the circularity worry brings out is a choice point for my event-type theory of episodic generic memories: (1) analyse knowledge of what it's like in dispositional terms, opting for epistemic generationism about memory, or (2) analyse knowledge of what it's like propositionally, committing to knowledge preservationism. In this paper I ultimately remain neutral on which approach is best, as both can account for the two core features of episodicity and generic representation.

5. Alternative approaches and their shortcomings

I have proposed an account of generic memories that captures their two pre-theoretic features, episodicity and generic representation. In this section, I will consider two alternative analyses which respectively drop, or explain away, one of these features, thereby reducing generic memories to a more familiar category of remembering.

On the first alternative account, when one remembers generically, one is really remembering a specific occasion but lacks subject-level access to the relevant temporal fact that would help one place it. Call this the *epistemic alternative*.⁶³ On the second alternative analysis, when one remembers generically, one is really remembering the particular *objects* that one encountered on multiple occasions in the past. Call this the *objectual alternative*. I will argue that both the epistemic alternative and the objectual alternative fail to capture something important about the phenomenon of generic memory.

⁶³ Because this alternative approach accounts for generic memories in terms of ignorance, it is 'epistemic' in the same sense as Williamson's theory of vagueness (1994), which accounts for the apparent vagueness of concepts such as 'heap' or 'bald' in terms of ordinary users' ignorance of the precise semantic boundaries of these concepts.

5.1 The epistemic alternative

One way to avoid positing a sui generis category of remembering is to analyse generic episodic memories as, roughly, unidentified singular episodic memories. What I remember in *The Commute*, according to the epistemic alternative, would be a specific day from my past on which I made that commute—I am just unable to identify which of the many possible candidate episodes I am remembering. My memory itself is specific, but because I lack epistemic access to the relevant piece of temporal information about the originating experience, I am unable to place it. From my epistemic perspective, then, I could be remembering any day or other on which I biked to campus during that period of my life. This gives my memory its generic feel. The epistemic alternative thus explains away generic representation, analyzing the felt generality in terms of the subject's epistemic relationship to her memory contents, rather than in terms of how those contents represent the past.

The epistemic alternative is attractive insofar as it can give a unified answer to the intentional object question for all episodic memories: every memory, if it successfully represents, is about a past episode. It will be particularly attractive to proponents of preservationism, since the idea of discretely retained memory traces could be leveraged to explain how every genuine memory uniquely picks out a past event even in the absence of subject-level access to the relevant causal chain. And, finally, we do sometimes episodically remember specific events while being unable to report on or identify the relevant facts about specifically where or when that event took place.

The main difficulty facing the present alternative is getting the phenomenology of generic memories right. How exactly can an epistemic feature account for felt generality? On a flatfooted version of the view, a memory has a generic feel when the subject lacks the appropriate piece of temporal information. But this can't be right: for example, one can remember a specific childhood

birthday party while forgetting how old one was at the time or can remember the first time one tried Ethiopian food while forgetting whether that was in college or early grad school, without its seeming to the person that for all they know they might be recalling any occasion or other that fits the general description. (For the strongest version of the counterexample, let's assume it was also the last time the person ate Ethiopian food.) Indeed, the fact there is something it is like to be unable to place one's memory in context suggests that, far from resulting in a feeling of generality, missing the right piece of temporal information can reinforce the recognition that one is remembering a specific occasion.

A more sophisticated version of the epistemic theory might account for felt generality differently. What produces the feeling of generality is not the mere absence of information, but the inability to discriminate between one event memory and all the other similar candidates. Recalling my commute, I might think to myself, "For all I know, I'm remembering a Monday ride and not a Tuesday ride. But since I can't tell the difference, there's effectively no fact of the matter about what day I'm remembering." On this account, the reason that the individuation question ("What time are you thinking of specifically?") doesn't seem to apply is because there is no way for the subject of the memory to answer it.

Even from the subject's perspective, there is an internal tension in this line of thinking. Is there a specific remembered event she thinks she could in principle discriminate (even though she has been unsuccessful), or is there not? Can she just not tell the difference between a Monday and a Tuesday memory, or is there really no difference to be tracked? These questions suggest that there are two distinct attitudes that can result from the epistemic task of identifying one's memory. In some cases, the unidentified memory retains an open-question feel—the feeling that one would be able to place it, if only one remembered in more detail, attended more closely, or recruited the assistance of someone else's help (e.g. a knowledgeable friend or psychoanalyst). In other cases, the unidentified

memory is accompanied not so much by the expectation that one will never find out but rather the certainty that there is nothing to find out anyway.

To illustrate the difference, suppose I recall wiping out while skiing down my favourite mountain. I remember its being late in the day and having tired legs, spotting the icy patch too late, and the sensations of skidding, falling, then trudging back up the slope to retrieve my lost ski. I can't recall what run that icy patch was on, whether I called it a day afterwards, whether this happened during a pre- or post-pandemic trip, and countless other details. Suppose I've had many such wipe-outs under similar conditions, making the individuation task difficult.

I can recall these contents in one of two modes. In the first mode, the contents present as imprecise or gappy; I cast my mind back in search of a crucial detail or a clue which would trigger an identifying inference (e.g. were those rental skis, meaning it was pre-pandemic? Did the ice catch me off guard, meaning conditions were generally good that day?). Even if my efforts are unsuccessful, I remain sure that I have in mind one unique occasion and can't help wondering which—even if I recognize no new information is forthcoming. My mnemonic inquiry is simultaneously open and frustrated.

In the second mode of recollection, by contrast, there is no similar sense of incompleteness or pending discovery. Because I have skied so often, fallen so often, and moreover imagined myself falling while skiing so often, I may recognize that my memory imagery is linked to a multitude of originating experiences, real and imagined. When I adopt this perspective, my memory doesn't call out for further inquiry into its experiential origins. I don't expect that introspecting my memory imagery will illuminate anything about just one event (although it may facilitate the retrieval of a separate, specific memory).

Perhaps this felt generality is due to my having already tried and failed to individuate my memory. It's plausible, for example, that the more I try to uncover further details, the less certain I

am that I am not fabricating other remembered elements, thereby undermining my memory's original felt specificity.⁶⁴ But perhaps my memory was always too vague or indeterminate to have felt specificity (because ski days are such a blur of activity, exertion, and sensory overload). Perhaps it features aspects that were knowingly filled in from other sources (such as my imagination, my skiing buddies, or my general knowledge of what happens when you fall). Or perhaps its contents were abstracted or blended from a number of thematically linked experiences.

Whatever the underlying process, felt generality includes a distinctive recollective attitude: I treat my memory's contents as a representational prototype for events of a certain description. Contra the epistemic hypothesis, the real reason the individuation question does not seem to apply to my memory is *not* that I cannot answer it—after all, in the former case I also cannot answer it, yet the referential inquiry remains open—but because my memory perspective is in fact not phenomenologically characterized as my thinking back to a single occasion.

If the foregoing distinction in recollective attitudes towards an 'unidentified memory' is legitimate, then the mistake of the epistemic alternative to generic memories is twofold. The relevant epistemic feature that the account appeals to—indiscriminability—is neither sufficient nor necessary for felt generality. It is not sufficient because one can fail to identify their memory even while the question of what, specifically, one is remembering remains open to the subject. It is not necessary because felt generality need not be the result a process beginning with an unidentified singular memory plus an individuation task. Although there is one plausible route from indiscriminability to felt generality, this explanation is not robust enough to capture the class of generic memories and, more generally, doesn't leave room for important phenomenological distinctions. The generic

⁶⁴ Even in such a case, I think there is a genuine transformation from what we would call an unidentified specific memory at t_1 to a generic memory at t_2 , during the progression of which details are lost, semanticized, reconsolidated, or reconstructed upon retrieval.

representation aspect of generic memories should therefore not be reduced to an epistemic phenomenon.

5.2 The objectual alternative

One might instead insist that the phenomenon of generic episodic memories can be accounted for within an existing alternative theoretical framework in the memory literature, one which already carves out conceptual space between remembering particular events and remembering general information by suggesting that subjects sometimes enjoy a more direct mnemonic relationship to the *objects*—people, places, things—that featured in their past experience.⁶⁵ Applying an object-centered approach to the present case, we get the *objectual alternative* to my analysis. The objectual alternative says that generic memories do not represent event types but rather the objects that were present across those event instances. In *The Commute*, I recall my bicycle, the roads I biked on, the landmarks I biked past, and so on. I do so by representing the perceptible properties of these objects, which I have abstracted from multiple past perceptual experiences. My memory represents the feel of my bike, the look of the river, the sounds of the traffic, and so on, without thereby representing any corresponding experience of *biking*, *seeing* the river, *hearing* the traffic, and so on. Since it does not pick out any originating experience, my memory is not temporally specific.

A core distinction underlying the objectual analysis is due to the work of Hoerl (2014), who defended the difference between remembering seeing an object—what he calls ‘event memory’—and remembering what an object looks like—what he calls ‘memory of looks’. His key example involves a

⁶⁵ See Debus 2007 for an argument that one can experientially recall an object without experientially recalling an event involving it, such as a past perception, because the latter but not the former memory involves representing the object “in a spatially involved way and at a determinate location” (p. 182). Debus (2008) later develops a relational account of experiential memory according to which one’s memory is partly constituted by the object(s) to which one is mnemonically related.

series of repeated visits to the Louvre to look at the Mona Lisa. He argues that he can recall what the Mona Lisa looks like without recalling any particular occasion of looking at the Mona Lisa (p. 352). Although the former ability causally depends on there having been repeated perceptual experiences of his representing the Mona Lisa, the contents of those experiences can be abstracted and entertained independently from the recollection of any one such experience. Look-memory is thus a non-episodic form of remembering objects, via perceptual information retained from past episodes involving those objects.

A theory of objectual memory has been defended more recently by Openshaw (2022). Drawing on Hoerl's distinction, Openshaw argues that one can remember objects by representing their past perceptible qualities: "to experientially recall *an object* is to activate knowledge of what it was in general like to perceptually encounter that object; e.g., of what it was like to see it" (p. 12). Openshaw's target of criticism is "eventism," which is the thesis that experientially remembering an object is only possible in virtue of remembering a previous event in which that object was encountered. On Openshaw's view, objectual memory is like episodic memory to the extent that it involves the production of phenomenally rich imagery. But unlike episodic memory, which represents objects and events from an egocentric point of view, objectual memory does not involve the adoption of a point of view within the scene. One can recall an object's past perceptible qualities directly, that is, without the spatial-perspectival characteristics that mark out a past experience. For example, I can visualize my brother's face, independently from any determinate spatiotemporal location involving me, looking at him. In so doing, I activate knowledge of what my brother looks like—a recollective act which causally depends on, but does not represent, past perceptual events.

Both Hoerl's look-memories and Openshaw's objectual memories can be helpfully understood as a limit case of *observer memories*: episodic memories in which the remembered objects are represented as from a perspective different from the one that the remembering subject occupied at the time of her

past experience. But rather than featuring an ‘inauthentic’ point of view, as observer memories do, objectual memories feature no specific point of view. And since there is no specific point of view, there is no temporal specificity inherited from a past perceptual event where that point of view was formerly occupied.⁶⁶ The objectual alternative therefore appears to capture the generic representation aspect of generic memories. For what is remembered is not a temporally local *occurrence*, like an event or a perceptual experience, but an enduring object; and the way the relevant object is represented in memory is via contents which have been abstracted from multiple perceptual encounters.

For observer memories at the limit, of course, episodicty drops out alongside perspectival contents. For if my mnemonic relationship to the objects of my past perceptual encounters is perspective-free, my act of remembering will not involve any sort of mental projection to a past point of view. It will thus be difficult to explain how my memory could feature any replay phenomenology, or the sense that what I am remembering is something that I experienced on (multiple) previous occasions.⁶⁷ The objectual alternative thus accounts for genericity by eliminating episodicty.

I am happy to grant that aperspectival experiential memories of objects make up a distinctive and important category of remembering in its own right. But as a theoretical framework with which to analyze memories such as *The Commute*, the objectual alternative is too restrictive. For in *The*

⁶⁶ This isn’t to suggest that remembering from a point of view *different* from the one actually occupied by the subject at the time of the past experience will render her memory temporally afloat from the relevant perceptual event. Against the idea that field to observer perspective switches are epistemically distorting, McCarroll and Sutton (2017) argue that non-egocentric visual information is available to perception at the time of encoding, which explains why it enters (veridically) into the construction of the memory. This same fact could potentially be leveraged as a reason to think observer memories are time-indexed in the appropriate way.

⁶⁷ A distinct but nearby worry is that the elimination of anything perspectival from the mnemonic relation between subject and object will make it difficult to account for any recollection phenomenology on the part of the subject. As Hoerl and Openshaw both recognize, the reproduction of retained perceptual information is not exclusive to mental acts of remembering. When I merely *imagine* my brother’s face, for example, I am similarly conjuring visual imagery that appropriately depicts it, imagery which was likewise causally derived from past perceptual encounters. On the objectual alternative so construed, nothing about either the contents or mode of representation would distinguish an act of purely objectual recollection from an act of experiential imagining. Indeed, this observation generates Hoerl’s concern about whether remembering looks should strictly speaking count as a kind of *memory* and leads him to describe it as “the epistemic face of visual imagination” (2014, p. 353).

Commute, I not only represent the general perceptible qualities of objects such as my bike, the roads, and the river. I also represent myself biking down the roads with the river to my left. That is, the perceptual knowledge I have activated in memory is presented not in a detached, viewer-free manner but gets worked into a scene that unfolds in my mind's eye as though from my past experiential perspective. In order to capture this mode of presentation, it is better to treat my knowledge in the first place as including perspectival elements. The notion of knowledge of what it's like to experience an event, from my account, is thus a more suitable candidate for what a generic episodic memory expresses, than the notion of knowledge of what an object looks like independent from any vantage point, per the objectual alternative.

In fact, an object-oriented approach is the wrong tool for the job not only because episodicity falls out of the analysis, but because of why it falls out: genericity was located in the wrong place. The kind of genericity that the objectual alternative generates is genericity over different *perspectives* on an object, not genericity over different *episodes*. But generic episodic memories can feature specific perspectives while being about unspecific occasions; assimilating the phenomenon to the limit case of observer memories will thus not target the right combination of features.

6. Conclusion

Remembering one's past generically is a rather ordinary feature of mental life that has nevertheless received little attention in the philosophical literature on memory. One reason for this lacuna is that the episodic–semantic memory distinction makes it initially difficult to see how to divorce episodic modes of representation from representational uniqueness. Taxonomy aside, a deeper problem is that it is hard to see how generic episodic memories could be both accurate and authentic at the same time.

Motivated by this challenge, I proposed an account of generic episodic memories on which they represent event *types*, but via experiential contents. Genericity over episodes is best explained in terms of the subject's knowledge of what it was like to experience events of that type.

Chapter Three

Perspectival Forgetting

Abstract: There are certain episodes from our past which we can replay with vivid detail, but which nevertheless remain distant from us in one crucial respect: we have lost access to the emotional or evaluative perspective from which we experienced those episodes. For example, you may be unable to recall enjoying a genre of music you now dislike, or unable to recall feeling hopeful after your expectations have been dashed. Such forgetting is all-too-familiar yet widely overlooked in recent work on memory. In this paper, I introduce the phenomenon, defend its status as a genuine kind of forgetting, and explain its wide-ranging implications to central debates about empathy and the powers of imagination.

1. A familiar kind of forgetting

Consider the following cases:

Lost Hope. You were recently hopeful that you were going to receive a promotion at work. You felt confident in your performance and, moreover, received encouraging but ultimately misleading testimony from one of your colleagues supposedly in the know. The morning of the big announcement, you shared your excitement with your best friend. But after the disappointing news sinks in, you have difficulty remembering feeling hopeful or believing you would get the promotion, despite your otherwise painfully vivid recollection of confiding in your friend.

Teenage Quarrel. As a college student, you had a particularly heated quarrel with your parents when you informed them you were switching majors from biology to philosophy. You felt misunderstood, aggrieved, and stubbornly in the right about your decision. Looking back on that event, you can remember all the aspects of your parents' reaction that infuriated you at the time. But having now experienced the challenges of the career choice and having been on the other side of advising conversations, you can no longer recall your former feelings in any great strength, despite knowing exactly what had prompted them at the time.

I take it these and other like experiences are familiar ways of relating to one's past self, particularly following a change in perspective—a shift that can happen quite suddenly, as in *Lost Hope*, or more gradually, as in *Teenage Quarrel*. The important feature to highlight for the present discussion is an element of the recollective experience, that is, the subject's memory following the perspective change. The memory is marked by what I will call a loss of access to one's past point of view, insofar as the way she originally experienced some episode—such as the hope she felt in the days leading up to the promotion announcement, or the anger she felt during the quarrel—is no longer available to her when she thinks back to that episode.

I claim that these cases exhibit a kind of *forgetting*. But what, exactly, is forgotten? These cases do not directly map on to paradigmatic forms of forgetting, such as forgetting how to do something (e.g. the method of long division), forgetting pieces of information (e.g. the name of a childhood friend), or forgetting events from one's personal past (e.g. eating breakfast exactly one month ago).⁶⁸

To see why not, let's take these hypotheses in turn.

⁶⁸ These three forms of forgetting are paradigmatic in the sense that the forgotten objects correspond to those of the three memory systems: an ability or skill (procedural memory), a piece of learned information (semantic memory), or a previously experienced event (episodic memory). See Frise (2017) for a helpful overview.

First, we can make sense of *Lost Hope* without assuming the subject has lost the relevant suite of abilities related to feeling hopeful, in general. She may have felt hopeful about something else only yesterday, an emotional episode which she can recall easily; she may be able to imagine feeling hopeful about something in the future. This is consistent with her having difficulty inhabiting that feeling in connection with her particular memory of the fateful announcement in *Lost Hope*.⁶⁹

Second, *Lost Hope* does not involve the loss of information, nor loss of access to information, that characterizes the kind of difficulty involved in recalling facts. For, in this case, the subject can recall *that* she felt hopeful, along with her former reasons for which she held that attitude, as well as the actions she took in light of that attitude. She can report all of her beliefs on these matters—including, notably, her beliefs about her former mental states—and thereby supply an objectively complete description about the past episode.

Third, typical varieties of episodic forgetting feature total loss of the experienced event (where whole episodes vanish into the past), partial loss (vague or incomplete representations of events), or loss of context (being unable to identify characters or places).⁷⁰ But in *Lost Hope*, the course of events remains vivid enough and otherwise complete, and there is no first-personal difficulty placing the events within the chronology of one's life.⁷¹

Nevertheless, something is missing in both *Lost Hope* and *Teenage Quarrel* which is typically present in episodic memories. Episodic remembering has a distinct representational format: it mirrors

⁶⁹ Obviously, there's a sense in which the subject is 'unable' to do something that would enable her to successfully remember in *Lost Hope*. I elaborate on the nature of the underlying difficulty later on in the paper. The important point at present is that an ability is not *itself* the thing that's forgotten, which is the why treating this case as an instance of procedural forgetting (like lost know-how, vanishing skills, dead habits, and so on) is not theoretically fruitful.

⁷⁰ McCarroll (2020); Michaelian (2011; 2023).

⁷¹ The same three observations about the object of forgetting hold for *Teenage Quarrel*. Briefly: The subject has not lost the ability to feel angry, nor has she forgotten why she was angry, nor does the problem consist in representational gappiness or unclarity about its location in her life history. The form of forgetting featured in the example is thus none of the usual kinds discussed in the literature: procedural, semantic, and episodic.

the perspective from which one perceptually experiences an event. What this means is that, paradigmatically, when one has an episodic memory, the information about the past episode is presented ‘from the inside’. For example, in recalling your colleague’s confident testimony, you not only remember that he said such-and-such but do so by representing the sound of his voice; in recalling the morning of the announcement you remember what you were wearing by visualizing the outfit.

Another way to gloss episodic memories is that they represent past events as though one is ‘re-experiencing’ them. That is, the subjective character of a successful episodic memory is qualitatively similar to the subjective character of the original experience. For example, visually recalling a blue car is qualitatively similar to perceiving a blue car, whereas merely remembering the fact that you saw a blue car (in the absence of visual imagery) is not.⁷²

A fortiori, successful episodic memories in which a felt emotion or attitude was a core component of one’s past experience will share phenomenal similarities to the past emotional experience. Remembering feeling happy yesterday involves—when all goes well—representing what that emotion felt like ‘from the inside’, that is, replaying yesterday’s events ‘as though’ I am currently happily experiencing them. This internal aspect of the remembered emotion is missing from both *Lost Hope* and *Teenage Quarrel*: the subject cannot recall things from the point of view from which that emotion was felt.

⁷² There is a parallel distinction for two forms of imagining. Sensory imagining (visualization being a familiar modality) has a similar quality to perceiving, whereas propositional imagining is—like supposing or conditionally believing—much less rich in its phenomenology.

2. The ‘nothing is forgotten’ objection

As mentioned above, episodic memories represent past events as though one were experiencing them again. Such talk of ‘replaying experiences’, or ‘feeling from the inside’, naturally lend themselves to an interpretation of remembering emotions more generally on which remembering some feeling *F* involves (in the good case) literally feeling *F* again while remembering. For example, it’s intuitively plausible that recalling a sad experience, such as the death of one’s pet, will typically be accompanied by feeling presently sad about the loss.

But this observation can suggest that emotions are not objects of memory so much as responses that a memory can generate. For instance, Dorothea Debus (2007) has defended the view that past emotions are, strictly speaking, never remembered, but only felt presently anew, upon remembering an event that once elicited them. Emotional memories, on her view, are analyzed as a conjunctive mental state consisting, on the one hand, of the event-memory, and on the other, of an affective response directed at the objects of the memory (which she calls “autobiographical past-directed emotions”).

Notice how, on a view like Debus’s, there is nothing in fact *forgotten* in cases like *Lost Hope* or *Teenage Quarrel*. The subjects of the memory simply no longer feel what they used to feel. And—Debus might add—the absence of the emotion shouldn’t be surprising. After all, the objects of the original emotion are located in the past, and it is no longer appropriate to feel the corresponding emotion presently. In *Lost Hope*, for example, with the possibility of the promotion foreclosed on, there is nothing to now feel hopeful about; and in *Teenage Quarrel*, given your new perspective on your career choice, there is no reason to still feel angry at your parents for their earlier reaction.

But the Debus model—on which remembering a feeling amounts to feeling that way again—doesn’t generalize well. This is because different emotional attitudes have different temporal structures

built in. Some, such as grief, are apt to persist diachronically, insofar as the object of the past loss continues to be lost. Here it is plausible that accessing your earlier grief, and feeling grief presently, are really the same way of responding to the past loss (though they may vary in their intensity, and so on.).

But many other emotions, such as surprise, resist this analysis. Suppose yesterday I heard what I thought was an infant crying outside my bedroom window, which, to my surprise, turned out to be a baby deer. Today, I'm reminiscing about that experience. On the present Debus-inspired analysis, it seems we are forced to say one of two things: Either (i) I can at most remember the *fact* that I was surprised, but not what that felt like 'from the inside,' or (ii) I can access that feeling from the inside by effectively feeling surprised again.

Both options are problematic. The first (i) imposes ad hoc constraints on the possible contents of my memory: that it includes, for instance, features like the visual appearance of the baby deer from my bedroom window (what it looked like to me), but it *excludes* my affective response to the deer (what it felt like to realize what I was seeing). Yet both aspects were part of my phenomenal experience at the time of the event, and thus both seem equally amenable to the experiential mode of representation that characterizes episodic memories.

The second horn (ii) implicates me in a kind of irrationality at the time of recollection. Surprise is fitting to feel in response to, roughly, thwarted expectations. Yet sometime shortly after the event, I had given up or updated those expectations: my new hypothesis regarding the source of crying sounds outside my window was that a baby deer, not an infant human, is present. Re-feeling surprise from my current perspective on the past object makes no sense.

The Debus alternative operates within an overly restrictive theoretical space: a subject can relate to her former attitudes either purely externally (as a mere fact about her mental state history) or purely internally (fully being in that same mental state again). The limitations of this model are clear

enough to set aside the objection that nothing is forgotten in the cases of *Lost Hope* and *Teenage Quarrel*. More positively, it highlights the need for the right kind of conceptual tools with which to understand the relationship between memories and emotions. As I will go on to explain, the familiar notion of imagining from a *point of view* or *perspective* is just what's needed to create space between the binary options of facts and feelings.

3. Imagining other perspectives

Here are some kinds of questions we invoke imagination to help answer. Will I enjoy my upcoming backpacking trip? How does my friend feel while public speaking? Had I skipped my swimming session this morning, would I have felt more relaxed now?

What these various imaginative projects have in common is they involve the representation of some scenario—real or counterfactual, present or future—from the vantage point of the subject experiencing it. I imagine future me attempting to be self-sufficient in the wilderness for days on end, and simulate experiences of exhaustion. I imagine how the crowd looks to my confident friend up on stage, and simulate feelings of exhilaration. I imagine a counterfactual version of my present self, physically restless, and simulate feelings of regret. In each case, I model an aspect of the imagined subject's experience, such as what they are sensing, feeling, or thinking.

In none of these scenarios am I actually undergoing the imagined experience. Nor am I simply engaging in suppositional reasoning about what kinds of mental states are predicted to obtain in circumstances of a certain description. I am doing a third thing altogether: representing the relevant attitudes from *within the scope* of the imagined perspective on the scenario.

My suggestion is to understand the structure of a mnemonic relationship to a previously held attitude in the same way. When you remember feeling F, you are imagining being that person (namely, you at an earlier time) feeling F. This doesn't require you to presently feel F in a way that would crowd out your current feelings (which may not involve F at all). But it does require you to do the imaginative work of reconstructing the point of view from which feeling F makes sense. And, for reasons I discuss later in this paper, it's not always possible to do so successfully—hence the possibility of forgetting.⁷³

That, in very rough outline, is the framework for experiential memory I will defend. On my view, forgetting what it was like to feel some emotion is, at bottom, a failure of *imagination*. In the remainder of this paper, I will argue that this view is useful, not only for making sense of the phenomena I have just introduced, but also for making progress within central debates concerning the relation between imagination and memory.

4. The puzzle of perspective-taking

Let's take stock. I have motivated an imagination-based analysis of certain cases of experiential forgetting. These cases—including *Lost Hope* and *Teenage Quarrel*—are marked by the absence of a feature that is paradigmatically present in successful cases of remembering attitudes: the internal

⁷³ Talk of perspectives in memory helps us make sense of other familiar phenomena involving the temporal displacement of an emotion, such as nostalgia for the past. States of nostalgia are paradigmatically mixed-valenced ("bittersweet"). On the one hand, the remembering subject is recalling a positive feeling, such as a happy childhood moment. On the other hand, the subject is also responding to the temporal displacement of the positive feeling, for instance wishing that happiness was not located in the past. That response is made possible by the fact that one's current perspective (less than fully happy) and one's former perspective (happy) are recognizably distinct.

Cf. De Brigard's (2017) theoretical framework for the contents of mental simulation. He argues that distinguishing (i) the feeling that a subject simulates from (ii) what the subject feels *while simulating* is necessary for making sense of many counterfactual attitudes, including regret and relief.

intelligibility, or phenomenal force, of the attitude. When this feature is absent from a subject's episodic memory, she forgets what it was like to have experienced the episode in the way that she did.

Moreover, forgetting what an experience was like is compatible with otherwise remembering the full context of the past experience, including the events that produced the experience or the reasons for which she held her former attitudes. The forgetting in question is thus not attributable to the usual culprits, such as information loss, trace decay, or mis-cueing during retrieval.⁷⁴ The problem involves imagining one's previous point of view.

I haven't yet given a positive account of what the difficulty amounts to. I have drawn attention to a potentially relevant feature: the cases are marked by a change in perspective between the time of the initial experience and the time of the recollection. Plausibly, something about the subject's *now* seeing things differently makes it difficult for her to remember things in the way she did. But I have also suggested that recalling a former experience is structurally similar to imagining someone else's experience, in that—paradigmatically—one imagines what things are like within the scope of the represented point of view without those contents conflicting with one's actual, present point of view. Together, these two suggestions generate a question: why should perspective-taking fail in *Lost Hope* and *Teenage Quarrel*?

This question interacts with a more general debate about the powers of imagination, and the conditions under which it can succeed in crossing an experiential gap. So, in answering our guiding question, I will also discuss the views of the two of the main camps in the imagination literature, the 'scaffolders' and the 'transformers.'

But my reason for introducing these additional interlocutors is also substantive. As I will shortly explain, our guiding question also gives rise to a puzzle. The puzzle is to explain why perspective-taking fails in *Lost Hope* and *Teenage Quarrel*, while at the same time allowing that our

⁷⁴ Cf. Schachter (2001).

personal past is a resource we can draw on in order to better imagine the experience of other people. This puzzle constrains the set of satisfactory answers to our guiding question. Scaffolders and transformers both founder against the puzzle, in different ways. Their attempts to address the puzzle is nevertheless instructive and will set the stage for a better account of experiential forgetting.

Here is a familiar description of an ordinary perspective-taking task. Suppose a friend of yours, a younger graduate student, is telling you about how nervous they feel about their first time teaching. You want to be able to empathize with the student—not merely by listening to what they’re going through, but by actively imagining what things are like for them. As a now seasoned veteran, you no longer get nervous before teaching. But you used to, back when you yourself were in grad school. One thing you can do, therefore, is draw on your past experiences of pre-teaching nerves in order to imagine what things are like for your friend. Intuitively, the act of empathizing will be successful insofar as the perspective you’re scaffolding from, and the perspective you’re aiming to imagine, are similar in the relevant respects (roughly: early year grad student wanting to perform well at a daunting new task).

On its own, the scaffolding-from-experience explanation is satisfying enough. But the explanatory adequacy quickly vanishes when we attempt to treat past-directed and other-directed imagining symmetrically. After all, the relevant respect in which your friend differs from you (they get nervous before teaching; you don’t) also characterizes your relation to your past self (you used to get nervous before teaching; now you don’t). If there was a legitimate question about how you can imagine your friend’s perspective, the same question arises about how you can imagine your past perspective. And, of course, we can’t appeal to the idea of drawing on past experiences for imaginative resources, since that very act is now the thing requiring explanation. Call this the *puzzle of perspective-taking*.

The circularity worry for scaffolding is a bad result for theorists who are optimistic about imagination's epistemic uses.⁷⁵ The idea of scaffolding from past experience is appealed to not only in connection with empathy, but also shows up in discussions of prospective imagining. Amy Kind (2020), most notably, defends the notion of scaffolding to explain how we can imagine the perspectives of, and make decisions on behalf of, our future selves. She claims that we can successfully imagine undergoing a new kind of experience by employing the experiential knowledge acquired from similar enough past experiences. For example, one can scaffold from one's past experiences of being a devoted pet owner—its various joys, cares, and burdens—to imagine what it would be like to become a parent for the first time. Imaginative projects of this kind would face similar circularity worries. How could our past experiences help us imagine future ones, if accessing a past experience is just one such imaginative project?

To rescue scaffolding from the circularity worry, then, its proponents are likely to deny one of the observations motivating the puzzle: the observation that I stand in the same epistemic relationship to my past perspective as I do to a future or another person's perspective. They will insist on the pre-theoretical observation that I can draw on my past experiences to successfully empathize or imagine my future, and so deny that I need to use my imagination to access my past experiences. More generally, then, they will deny that imaginative perspective-taking should be given symmetrical treatment across the three contexts of remembering, empathizing, and future imagining.

Indeed, the epistemological principle behind scaffolding arguably already involves an implicit rejection of symmetry. The assumption behind the possibility of scaffolding is that previous experiences are—insofar as they are remembered—simply *available* as resources for one's various imaginative projects. In empathizing with the younger grad student, I simply draw on some stored experiences of mine. The conception of memory operative in the background here is fundamentally

⁷⁵ Williamson (2016); Arcangeli & Dokic (2018); Kind (2020; 2021); Langkau (2021); Bailey (2022)

preservationist: memory stores contents from previous experiences, in traces, ready to be activated or retrieved in the service of imagination. On this picture, memory's role is to *retain* the material of our experiential resources, whereas imagination's role is to *use* these experiential resources as building blocks for perspective-taking.⁷⁶

Resolving the puzzle in this way, however, exposes a residual problem. By treating past experiences as raw material for imaginative projects, we are no longer able to shed light on the *difficulties* of remembering that I claim characterize cases like *Lost Hope* and *Teenage Quarrel*. After all, what is distinctive about these cases is that they involve, not the inability to retrieve stored information, but rather the inability to re-inhabit a past perspective. In other words, it is hard to see how forgetting—understood as a kind of imaginative failure, rather than, say, a memory storage problem or loss of information—could occur, if memory and imagination have the roles that scaffolders assign to them.

Moreover, the idea of scaffolding from one point of view to another has its detractors, as well. Many philosophers think it is not so easy to imagine one's way into a different perspective as the optimistic picture above suggests.⁷⁷ Another available response to the puzzle, therefore, is to assign less weight to the observation about easy empathy, more weight to the idea of symmetrical treatment, and thus insist that—regardless of its temporal directionality—acts of perspective-taking across large enough experiential divides are likely to fail.

For instance, Laurie Paul (2014, 2017) argues that one cannot successfully imagine experiencing things from a future perspective in cases where the future perspective is structured by 'core values' not currently shared by the imaginer. Scaffolding across a big experiential divide fails

⁷⁶ One way of understanding what generates the original puzzle is the clash between two conceptions of the relationship between remembering and imagining: preservationism and constructivism. Preservationists adopt some version of the storage-retrieval model and the idea of causal traces (Bernecker 2010). Constructivists treat the notion of retrieval as itself an act of imagination and re-conceive the notion of causal traces (Michaelian 2016; De Brigard 2014).

⁷⁷ D'Arms (2000); Goldie (2005, 2011); Paul (2014, 2017); Bailey (2022, 2023a, 2023b)

when the nature of the target experience is such that it would transform the very ways that the imaginer responds to the world—what they care about, their emotional capacities, and so on.

Olivia Bailey (2022, 2023a) develops a similar thought in contexts of empathy. On her view, if the empathizer’s ‘emotional sensibilities’ (viz. their characteristic ways of responding to perceived value) do not match the sensibilities of the person they are attempting to empathize with, they may fail to understand the other’s experience. That’s because (i) imaginatively reproducing another person’s feeling depends on sharing the relevant evaluative response to an object or situation—such as felt anger at a perceived injustice—and (ii) the responses that are possible for any individual depends on their existing sensibilities—such as a tendency to experience certain action types as unjust.⁷⁸

For instance, on Bailey’s picture, if I am the kind of person who finds country music distasteful, I will be unable to empathize with your enjoyment of it. You pointing out its good-making features will not help me feel what you feel, if those features do not also strike me as aesthetically valuable. And my attempts to scaffold from my felt appreciation of another genre of music will not help me either, since I cannot simply re-attach the phenomenal aspect of that feeling to an object I cannot view as worthy of it (2022).⁷⁹

Here’s how the same line of reasoning would apply to the case featured in the puzzle. Let’s say you no longer experience undergraduate students in your classes as holding power over your career,

⁷⁸ On this point Bailey (2023a) agrees with Goldie (2011), who argues that a braggart will be unable to empathize with a humble person. Even if the braggart tries to imaginatively adjust for their difference in dispositions, there is a crucial difference between (i) reasoning from a character trait to a hypothetical attitude and (ii) representing the attitude ‘from the inside’, that is, as a first-order response to the world.

⁷⁹ Bailey’s target here is Langkau (2021), who suggests that phenomenal feelings of appreciation or enjoyment can be leveraged to imaginatively cross differences in taste.

For an interesting alternative, see D’Arms (2000) for a defense of *non-imaginative* empathy whereby one ‘catches’ the other person’s feeling first, and arrives at the corresponding evaluative judgment second. Like other skeptics of imaginative perspective-taking, D’Arms believes there is a risk of getting trapped in one’s current perspective when trying to imagine the perspective of someone else. But he uses this observation to argue that picking up the phenomenal character of the target attitude first, through a subconscious process like ‘emotional contagion,’ is ultimately a better route to shared understanding than, say, deliberate imagining or evaluative reasoning. For example, you might pick up on a friend’s anger during a dinner party, without knowing what caused it, which you can then use as a lens onto the world in order to uncover slights or insults you hadn’t originally experienced as such.

nor relate to your role as an instructor as a character you are obligated to act out. You now have a very different kind of response to the same features of the classroom that *previously* struck you as nerve-racking. Of course, you can remember what those features were. And perhaps you now feel nervous in other kinds of situations. But neither of these things, on the Paul-Bailey line, will help you imaginatively occupy—or reinhabit—the point of view from which that particular nervousness is felt. That’s why both empathizing and remembering are epistemically fraught.

Notice that a commitment to symmetry across imaginative contexts falls out of the principled pessimism about scaffolding. As far as access to a different perspective is concerned, it doesn’t matter, in and of itself, whether you used to occupy that perspective. What matters instead is whether your current view of things permits the range of evaluative responses to the world that would allow you to represent the target experience from the inside. A past experience can thus in principle be equally as ‘alien’ to a person as a possible future one.⁸⁰

The anti-scaffolding picture thus seems to offer a ready-made explanation of the phenomenon of forgetting what something is like. When the experiences a person undergoes substantially changes their view of things, such that the resulting evaluative outlook is incompatible with the initial one, the new outlook not only replaces the old one, but imaginatively crowds it out. After not receiving the promotion, the subject of *Lost Hope* can no longer see the same signs as having been evidence for a good outcome; after experiencing the difficulties of the career choice, the subject of *Teenage Quarrel* now sees her parents’ disapproval as farsighted rather than a rejection of her independence.

The worry about resolving the puzzle as the transformers would—denying imaginative access to both past experiences and experiences of other people—is that the case against scaffolding will overgeneralize, predicting too many instances of forgetting. Intuitively, not every change in our values

⁸⁰ See Bailey 2023b for an argument that the same principles which tell against the possibility of imagining certain future experiences will also apply to remembering certain experiences.

or emotional sensibilities will undermine the ability to remember such past perspectives. We can, and do, remember feeling excited about things we no longer consider worthwhile. Indeed, certain emotional responses *towards* our personal past, such as shame or embarrassment, wouldn't be fully intelligible if our past experiences and emotions were accessible merely as bits of information about ourselves. We experience shame and embarrassment precisely because we can, at least to some extent, re-inhabit evaluative perspectives that now seem foolish to us.⁸¹ More generally, the transformers' response to the puzzle will leave them unable to distinguish between those evaluative shifts that do result in forgetting and those that don't.

What the puzzle of perspective-taking brings out, therefore, is that proponents and skeptics about imaginative scaffolding overshoot in different directions. The scaffolders overshoot by going in for a picture of imagining on which remembering past experiences is trivially easy. The transformers overshoot by adopting principles about perspective-taking which render remembering past experiences overly difficult. None of these approaches offer a systematic way to understand the connection between remembering and imagining.

Yet there is a grain of truth in each approach. The scaffolders' rejection of symmetry is not unmotivated, but rather pays due attention to the intuition that one's past perspective is often more familiar than either one's future perspective or that of another person. For unlike imagining an entirely new experience, when it comes to remembering one has actually undergone the experience one is attempting to represent. This fact should plausibly be assigned some weight in the calculus of epistemic access.

On the other hand, the pessimists are right to think access to one's personal past is not a given. It too requires an act of imagination, not unlike that involved in other imaginative projects concerning

⁸¹ Compare: the nature of shame or embarrassment that one feels when one is *told* about something one has done that one cannot remember (e.g. behavior while inebriated) versus the full-blooded version of the attitude one feels when one can relive, in memory, one's actions or experiences.

future selves and other people. And in all these cases, perspectival differences can be hard to overcome.

What we need, then, is an account of perspective-taking that can explain how it is that we can recruit our past perspectives in the service of empathizing with others, while at the same time allowing that we ourselves can forget what it was like to occupy a past perspective.

In the next section, I will begin to make room for such a notion of perspective-taking by providing a more detailed criticism of the transformers' skepticism about it. As I will explain further in section 6, it is not only possible but in fact crucial to the *success* of many imaginative projects that we maintain what I will call a 'bifocal' view of both our actual perspective and the one we are imagining. In section 7, I apply this lesson to the puzzle of perspective-taking, explaining how bifocality can allow us to make sense of cases like *Lost Hope* and *Teenage Quarrel* without generating implausible commitments.

5. Rescuing perspective-taking

To begin, let's take a closer look at why transformers think that scaffolding from one perspective to another is not possible.

Optimists about the imagination's epistemic powers, like Kind (2020), think something like experiential knowledge can be acquired through imaginative scaffolding. Scaffolding involves skillful processes of adding, subtracting, or otherwise modifying the material of our experiential resources. The underlying picture is the familiar empiricist one: if I have seen the color gold, and know what mountains look like, I can imaginatively combine their properties to form an image of a golden mountain.

Opposition to this picture is largely driven by the observation that experiences which engage our emotions and values—the kinds of experiences we care about in the contexts of empathy and decision-making—don’t permit atomic-level operations. That’s because the contents of emotional or evaluative experiences are intrinsically related to their objects. Caring *about* x, for instance, is not reducible to the conjunction of an attitude (caring) plus an object (x), because the attitude is a response to the particular combination of perceived properties of the object that make it worth caring about. That’s why you can’t factor out the caring aspect of a past experience and imaginatively attach it to a new object.

These claims about dissociability, if they’re right, prohibit scaffolding from the phenomenal end of an experience: the felt attitude or emotion. But that leaves an apparent workaround. Why not simply imagine the new *object* directly, with its distinctive properties, in order to access or simulate the relevant evaluative or emotional response that way? For example, my friend tells me about some recent difficulties at work. I imagine the situation she describes, including the careless actions of a colleague that led to everyone having to put in overtime. I arrive at an imagined feeling, which I can identify as frustration.

There are two reasons transformers seem to provide for why imagining the new object or scenario directly won’t work. One is that having the right object in view may *already* require seeing it in a certain evaluative light. Perhaps I have to antecedently accept my friend’s normative assessment of the situation at work to guide my imaginative project. I can’t simply represent to myself the low-level features of their situation, independently of their evaluative significance. In that case, it would turn out that imagining the object of another person’s experience is not a distinct first step in the process of coming to find out what that experience is like. I need to see things from within the scope

of the relevant emotional attitude to even succeed in representing the scenario that my friend is responding to.⁸²

The second reason is that, even if I can represent the imagined object independently from some evaluative or emotional point of view, all I will discover is how *I*, the present imaginer, am *already* disposed to feel in response. That is, the most I will succeed in doing is projecting my current patterns of thinking, feeling, and valuing, into different circumstances. I won't count as 'trying on' new ones.⁸³ So, if I am not the kind of person to get frustrated easily, I won't be able to imaginatively relate to my friend's work situation in the way that *she* does, because I will have a different evaluative responses to the low-level features of the situation that, for her, called out for frustration.⁸⁴

The common denominator behind both lines of reasoning is that one is effectively trapped in one's current perspective during acts of imagination. That is, one's current ways of seeing the world will either constrain how one *represents* imagined objects and scenarios, or how one imaginatively *responds* to these objects and scenarios. It's just not possible, in either case, for someone to intentionally

⁸² Cf. Bailey's (2023a) argument for why she cannot imagine experiencing hockey as an admirable sport. "Since my available experience of sport as admirable is wholly oriented around delicacy, it is hard to see how drawing on that experience will allow me to picture shoving and slamming as features that ground sporting admirability. I could, perhaps, come to picture hockey as delightful in my mind's eye by imaginatively shifting some of its lower-level features. I could imagine, for instance, that hockey players occasionally pause to artfully wave a ribbon, and in that case, my mental representation of hockey will be more fit to imaginatively draw together with my own experiences of sporting enthusiasm. But in that case, the imaginative picture I arrive at will quite dramatically fail to mirror your emotional apprehension of hockey as loveable. After all, ribbon-waving would absolutely ruin hockey's appeal for you" (p. 227).

⁸³ According to Peter Goldie (2011), imagining oneself in another person's circumstances, and imagining being the other person (in their circumstances), are two distinct cognitive acts. He calls the former "in his shoes imagining" and the latter "empathetic perspective-switching." The case he makes against empathy takes the form of a dilemma. Suppose you and I have different character traits. If I don't take those into account when I imagine being in your shoes, all I will succeed in representing is a hypothetical experience of mine. But if I do take those differences into account, I will have to imaginatively foreground facts about you (e.g. that you get frustrated easily whereas I do not) that don't show up from your own point of view, and thus misrepresent your experience that way. (The background thought here is that *you* don't reason your way from your own character trait whatever it is you're feeling, whereas I, the imaginer, would need to do so, to prevent my character traits from running interference. Yet as a result, I cannot end up with a truly first-person representation of your point of view.)

⁸⁴ This is the point Paul (2017) makes in the context of future imagining. If I am not the same kind of person I *would* be following the birth of a child, I won't succeed in imagining my *future* attitudes, but only my current attitudes about a future situation. "The future me, the person who I'll become, is first-personally inaccessible to me now" (p. 205).

step outside of their present modes of evaluation or affective dispositions. This is the core of the skepticism about imagining another point of view.

It remains possible, of course, to *change* one's perspective—e.g. by actually undergoing a new kind of experience (e.g. becoming a parent) or by being persuaded to arrive at different value judgments. But the notion of 'trying on' a point of view, where that consists in representing the world according to that point of view without actually adopting it, vanishes from their analysis.⁸⁵ The transformers' view is, at the end of the day, a rejection of the very idea of perspective-taking.

This seems to me to be a mistake. It underestimates the complexities of perspectival thinking. In trying to imagine what things are like for others, we often *bracket* rather than change aspects of our perspectives.

'Bracketing' is a familiar feature of many cognitive operations performed by the imagination. It is most obvious in cases where we bracket, for purposes of imagination, the contents of some of our beliefs. When we engage in counterfactual reasoning about how events could have unfolded instead, we suppose something contrary to what we know to be true, and reason about what follows within the scope of that supposition. And we do so, of course, without involving ourselves in any contradiction of epistemic commitments.

Sensory forms of imagination are no different. If I am trying to decide whether to paint my bedroom walls a new color, I can visualize what the walls would look like if they were light blue, even as I currently perceive them as off-white. I can imagine the imminent feeling of taking a cold shower

⁸⁵ Here is an example to demonstrate why the two exclusively available options seem to be (1) failure to imagine or (2) change of mind. Suppose I find a particular painting ugly and want to understand why someone else experiences it as beautiful. The other person points out some of its aesthetic qualities that I haven't focused on before, paired with an explanation of why, for them, these features make the painting beautiful. If, upon attending to these features in the manner suggested to me, the painting nevertheless 'leaves me cold', I haven't successfully imagined it as beautiful. And if I *can* see it in a new light, I have now—at least temporarily, if not stably—changed my view of the painting: I actually find it beautiful.

while walking home in thirty-degree Celsius heat. Imaginatively shifting features of a perceptual experience does not thereby change that experience.

Empathy, too, requires bracketing to even get underway. Say I'm planning a surprise party for my partner later this evening. Throughout the day, I act as though nothing special is going on. I know he believes that I forgot about his birthday and is upset at me about that. I can empathize with him, simulating the way he's feeling given the significance to him of my apparent omissions (even though I *know* his disappointment is ultimately baseless). That is, I bracket my own representation of the present situation, in order to project into his point of view, according to which my omissions are worth getting upset about.

Just like the imagined perspective does not come into conflict with my real one, the emotional response I simulate within the scope of that perspective does not come into conflict with my current emotion (anticipation, say). I do not, in other words, *actually* come to feel upset at myself in virtue of successfully empathizing with my partner. Successful perspective-taking, and the kind of understanding of the other person it affords, does not require what we might call real attitudinal change.

The transformers may acknowledge that *some* bracketing is involved in the case: the kind of bracketing that allows me to imaginatively represent my actions under a different description and thus as having a different normative significance from the one I know my actions to have. But for all the example shows, I am still operating within my overall *evaluative* perspective when I then proceed to simulate my partner's response. It's open for the transformers to insist that if the situation were as my partner believes, and if I were in his shoes, I too would actually be disposed to find it upsetting. So, perhaps my ability to successfully empathize still *depends* on something like a shared sensibility or value scheme.

But evaluative or emotional perspectives can be bracketed too, for they likewise are—or involve—mental representations. This is perhaps easiest to see if we set aside, for the moment, how we relate to other people’s emotions, and consider a familiar way of relating to our own.

Suppose I find myself experiencing something my friend is doing as infuriating. I realize I am angry. I reflect on this fact: I consider whether my experience presents my friend’s action in the right light. This might involve, for example, weighing the felt strength of the perceived slight against what I understand to be the reasons for my friend’s behavior. Upon reflection, (i) I determine that my response to my friend is unjustified, that what they’ve done does not warrant being angry about, and yet (ii) I continue to experience their behavior as infuriating. (Perhaps because they are hitting upon a pet peeve of mine, I can’t simply reason my way out of my habitual form of response.) The result is that I see the world through the lens of an evaluative attitude that I do not endorse. I am to that extent ‘alienated’ from my emotion.⁸⁶

Cases of alienated emotions reveal two things. The first is that first-personal experiences can be *bifocal*—that is, one can inhabit an emotional viewpoint (i.e., find it intelligible ‘from the inside’) without endorsing the view of the world it suggests (i.e., judge it, from the outside, to be incorrect). In other words, one can occupy a perspective while at the same time evaluating that very perspective from the point of view of another that conflicts with it. The second is that the possibility of bifocality depends on ‘stepping back’ from one’s original emotion, and relating to it *qua* representation—that is, as a take on the world that may, ultimately, be incorrect.⁸⁷ That’s just what bracketing one’s perspective consists in.

⁸⁶ See Boyle (2024, p. 312) for a discussion of alienated emotions.

⁸⁷ Cf. Korsgaard: “For our capacity to turn our attention on to our own mental activities is also a capacity to distance ourselves from them, and to call them into question. I perceive, and I find myself with a powerful impulse to believe. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn’t dominate me and now I have a problem. Shall I believe?” (p. 93) For an account of the structure of reflection that rejects the idea of backing up from one’s mental states, see Boyle (2024).

If bracketing opens up the possibility of reflective assessment, it certainly licenses *imagining* how things might look through a different lens. Indeed, I've suggested that reflective assessment already involves a kind of perspective-taking, because it requires adopting an evaluative stance *external* to one's first-order mental state (e.g. the question "Should I feel e ?" is asked from outside the scope of e , as it were).

6. Successful empathy

The point of highlighting bifocality in the context of self-reflection was to show how it is possible for someone to understand the intelligibility of a point of view without fully aligning oneself with it. What this means for empathy is that I can appreciate what someone else is feeling, 'from the inside' no less, without losing track of or giving up whatever it is *I* actually feel.

I now want to suggest next that bifocality also characterizes the psychological structure of *successful* empathy. On my view, it is not only psychologically possible for the imagining subject to operate with distinct (and sometimes conflicting) representations at once. Doing so is positively conducive to performing acts of perspective-taking *well*.

The bifocality featured in empathy has a different flavor from the 'alienated' kind of bifocality that can result from reflectively disavowing one's felt emotions. In the intrapersonal case—concerning me and my take on the world alone—seeing things one way from the inside of one's emotion, and another way from the outside, can be alienating insofar as both perspectives are centered on a single psychological subject. That's why there is a sense of internal conflict or instability. But when it comes

to imagining the perspective of someone else, the same psychological facts do not obtain, and so the mere fact of difference in perspective does not threaten subjective disunification.⁸⁸

Moreover, in the context of empathy, it's important to maintain a sense of differentiation between oneself and the person one is empathizing with. There are at least two ways in which a failure to keep this difference in view can undermine an empathetic project.

One way is by engaging in naive counterfactual *self*-imagining. For example:

The Breakup: Your friend is going through a difficult break-up with S and is telling you about all the things they miss about being with them. You try to empathize with your friend, by imagining what it must feel like to be in their situation. You can't help thinking that, if you were in their position, you would be relieved, since you never really liked S in the first place. You tell your friend, "I'm afraid I just can't understand how you feel. Isn't this for the best?"

The mistake in *The Breakup* is easy enough to spot. By not bracketing your own views about S and letting these override the reasons the friend is reporting for caring about S, you fail to do what the situation requires, namely, to see things from your friend's point of view. The friend's perspective has been crowded out by the evaluative stance of the empathizer.

But there's a second way empathy can go wrong in virtue of failing to maintain self-other differentiation:

⁸⁸ Other differences between the intrapersonal and interpersonal perspective-taking, which are worth highlighting but are not important for the present argument, concern the directionality and purpose of the act. In the intrapersonal case, I try to get 'outside' of my own attitude, relating to it as a mere representation, so that I can assess its grounds. In the interpersonal case, I try to get 'inside' of a different attitude, to understand why it makes sense as a first-order response to the world.

Breakup 2.0: Your friend is going through a difficult break-up with S and is telling you about all the things they miss about being with them. With some help from the testimony your friend is currently sharing with you, along with everything you know about your friend’s character, you recreate their perspective entirely. You imagine being on a first date with S, proposing to S, moving into a new home together, and the thought of this relationship being over brings you to tears. You feel moved to call S to try to fix things. You tell your friend, “I understand completely. This feels terrible!”⁸⁹

The mistake in *Breakup 2.0* is more difficult to spot, in part because the mistake that was present in the first case—failing to imaginatively adjust for differences between you and your friend—seem to be corrected for in the second. Indeed, cases like *Breakup 2.0* are often implicitly taken to be the epitome of successful empathy, insofar as the empathizer is undergoing a qualitatively identical version of the target’s experience.⁹⁰

But it seems rather to me that the empathizer has overcorrected for perspectival differences, by removing the very fact of difference from view. What’s missing from their imaginative project is a representation of the other person’s feeling as belonging to the other person. The intended subject of the imagining and recipient of empathy—my friend—has again dropped out of sight, only this time because I have fully adopted her attitudes as my own. My act of empathizing has slipped thus into

⁸⁹ This example is inspired by Canto V of *Inferno*: Dante, upon listening to the story of the weeping spirit of a sinner, is so overcome with empathetic feelings of pain and sorrow that he faints.

⁹⁰ This idealizing assumption is problematic for a number of reasons. First, it invites criticisms about empathy’s value as a practice. For instance, you might think it’s not a good outcome in *Breakup 2.0* that we end up with two people feeling bad, rather than, say, one person being helpfully responded to by another who can maintain a certain kind of emotional distance from the situation. Second, if success is conceived of largely in terms of the *outcome* of empathy—producing an experience with certain contents—we lose sight of why imagination-based empathy is any more valuable than its non-cognitive cousin, ‘emotional contagion’ (i.e. inadvertently catching the emotion of a person in your environment via the firing of mirror neurons). Third, it raises a distinctively ethical concern that imagining another person’s lived experience without due acknowledgement of its being theirs is a disrespectful or invasive way of relating to the other person. (See Kind 2021 for a discussion of ‘epistemic arrogance arguments’ against empathy.)

something closer to voyeuristic imagining or self-preoccupied fantasy.⁹¹ I have not so much imagined things from my friend's point of view as I have appropriated that point of view, by substituting my consciousness for theirs.

We can avoid both forms of imaginative mistakes—what we might call underprojection and overprojection into another point of view—by maintaining precisely the kind of bifocality to which I have drawing attention. I imagine your attitudes at once ‘from the inside’, such that they are intelligible to me, and ‘from the outside’, such that they point to you as the subject of my imagining.⁹² When I do so, I succeed in imagining what things are like for you.

7. Perspectival forgetting

We are finally in a position to make sense of our initial phenomenon: forgetting what something was like. The general view sketched in section one was that memories like *Lost Hope* and *Teenage Quarrel* exhibit a loss of subjective access to one's past point of view: the subject cannot recall, ‘from the inside’, what they felt during the past episode. I claimed that the recollective difficulty was not due to a loss of information, nor the mere fact that the subject, at present, no longer has reason to feel the way they used to, but something else: a failure of perspective-taking. These cases thus exemplify a distinctively imaginative form of forgetting.

⁹¹ The effect featured in the case is similar to the phenomenon investigated by psychologists under the name of ‘contagious distress’ or ‘over-arousal’. See Coplan (2011) for a discussion of the psychological and neurological differences between *self-oriented* and *other-oriented* varieties of perspectival imagining. These differences support the verdict I am offering here: *Breakup 2.0* results in a kind of self-oriented imagining.

⁹² For other helpful illustrations of the idea that imagining points of view necessarily involves indexing representations to subjects (real or hypothetical), see Velleman (1996) and Arcangeli & Dokic (2018).

As I will now explain, the same structural features of perspective-taking that are required to successfully imagine the experience of another person—viz. bifocality and bracketing—are required to successfully remember the experience of one’s past self. Forgetting what something was like can thus be explained in terms of failures of bracketing and loss of bifocality when imagining the past.

To begin, notice that episodic memory, like empathy, is bifocal in structure. This observation is not surprising. After all, a kind of bifocality shows up in descriptions of the characteristic *temporal* phenomenology of episodic memory, which combines talk of ‘replay’ and ‘pastness’. When I remember how I used to feel, I recall an event as though I am experiencing it (now) as I did at the time, while also representing it as belonging to the past.

I suggest thinking of this temporal structure along the lines of the perspectival structure of imagination more generally. When I succeed in remembering an emotional experience, two things are true of me. On the one hand, I represent a past event through the evaluative lens of the remembered emotion. It makes sense to me as something to feel; I feel its normative pull as a response to the events I am recalling. On the other hand, I maintain an external distance from that feeling, relating to it as a state I used to be in but (perhaps) am no longer in. The feeling is indexed to a perspective that is temporally displaced from me qua presently remembering subject.

One of the lessons of our discussion of empathy is that I can succeed in imaginatively representing feeling something, without thereby *actually* feeling it, i.e., in a way that would come into conflict with my present emotional state. The same is true for memory: contra Debus, I can ordinarily recall feeling sad yesterday while feeling perfectly happy today.⁹³

⁹³ As I explained in section two, Debus’s model equated remembered emotions to presently felt emotions about the past. We now have the tools to identify the same problems with that framework for memory in connection to empathy. Debus’s idealized conception of success for memory is equivalent to the idealized conception of success for empathy featured in *Breakup 2.0*. Just like in *Breakup 2.0*, there is a distinctive phenomenon studied by psychologists involving an equivalent loss of self-other differentiation for memory: “flashbulb memories”. Flashbulb memories, which are often linked to traumatic experiences, involve a representation of the remembered scenario in the absence of temporal displacement, that is, as though it were fully present to the subject. They result in the subject’s presently feeling an identical version of whatever they felt in the past.

Here too, our ability to *bracket* our present point of view is what makes it possible to intelligibly represent the world in a way that is distinct from how we actually take the world to be. Indeed, bracketing is what makes episodic remembering possible more generally: recalling something to mind that no longer obtains does not involve a person in any kind of contradiction. It's also why simply having *changed* your evaluative or emotional perspective, between the time of experience and the time of recollection, is not an in-principle barrier to successfully remembering what something was like.

If it is possible to remember emotional experiences after one's perspective have changed, why, then, does forgetting occur in cases like *Lost Hope* and *Teenage Quarrel*? The answer is that forgetting happens because bifocality has been lost. That is, because successful perspective-taking requires maintaining the appropriate bifocality, and because bifocality is made possible by bracketing, imaginative projects can fail when the subject doesn't bracket the right elements of her current point of view.

To illustrate, consider first an interpersonal version of the *Lost Hope* case. I am hopeful about receiving the upcoming promotion; you are on the committee, so you know that this isn't going to happen; you are also aware that I've received misleading evidence from another member. Knowing all these things, you can reasonably predict how disappointed I'm about to feel when I receive the news, because you can guess at my currently feeling very hopeful. Can you also *empathize* with my currently feeling hopeful? Only to the extent that you can quarantine the elements of your own perspective on the situation, from which vantage point there is no cause for hope. This will be more difficult for you to do if you are, say, anticipating, dreading, or otherwise preoccupied with how events are going to go once you deliver the bad news, since this foregrounds the reality of your vantage point, making it harder to achieve the requisite reflective distance.

If these explanations are on the right track, then it is plausibly even more difficult for the person who really ends up disappointed to bracket their perspective on the situation after the change.

For that person, the sense of loss is too dominant, too much a salient part of their experience at the time of the attempted recollection, to be successfully bracketed. The subject of *Lost Hope* thus cannot maintain the right temporal separation of attitudes when they try to imaginatively project themselves into their former point of view.

The explanation of forgetting in *Teenage Quarrel* has the same basic structure: the failure to bracket one's current perspective on the past at the time of recollection interferes with one's imaginative representation of their past experiences. But what makes the act of bracketing difficult might be different. Here's one plausible psychological explanation. Your career choice ended up being challenging in a number of (unforeseen by you) ways. It has the felt duration of ages. Your college years, during which you were moved by a sense of lighthearted freedom and defiant independence, feel not only distant from you in time, but as having happened to a different person. Teenage you thus lacks the psychological reality that you have, or that recent versions of you that are more continuous with present you, also have. This makes it harder for you to bracket your own perspective, since bracketing one's perspective is something we do precisely to access a perspective other than our own. In other words, the self-other differentiation is not robust enough to aid you in moving away from one perspective and towards the other.⁹⁴

⁹⁴Another way to put the thought is that your former way of feeling and seeing the world has become, for you, too theoretical—more of a description than a way of being that can be occupied. And yet, if you happened to encounter that same perspective in another person—say, a college student currently on the receiving end of paternalism who is sharing their anger with you—you might easily enough be able to imagine it again. In this way, sometimes empathy with another person can help one recall a past form or life, rather than the other way around.

8. Conclusion

Much has been written both about the role of imagination in remembering one's personal past, and about the kind of imaginative projection involved in empathizing with the emotions of another person. In this paper, I have focused on an overlooked phenomenon situated at the intersection of these two topics: what I call perspectival forgetting. This phenomenon helps bring into view a puzzle about how imaginative memory and empathy interact. Clearly, we can forget what it was like to affectively respond to events from our past as we did at the time. We can fail to 'empathize' with our past selves. And yet we can, and often do, recruit memories of past experiences to help us better empathize with other people. My framework of bifocality for perspective-taking makes progress on this puzzle by offering a unified analysis of the difficulties involved in empathy and emotional memory, as well as of what it takes for either of these imaginative projects to succeed. On my view, successfully empathizing with another person, or remembering what it was like to feel an emotion one no longer feels, involves maintaining the relevant self-other differentiation by bracketing aspects of one's actual evaluative perspective. We fail to remember our former attitudes from the inside when we are unable, for a variety of reasons, to bracket our current point of view at the time of remembering. Forgetting what something was like, in other words, is a failure of imagination.

Bibliography

- Abelson, R. (1981). Psychological Status of the Script Concept. *American Psychologist* 36 (7): 715-729.
- Addis, Donna R. (2020). Mental Time Travel? A Neurocognitive Model of Event Simulation. *Review of Philosophy and Psychology* 11 (2): 233-259.
- Andonovski, N. (2020). Singularism about Episodic Memory. *Review of Philosophy and Psychology* 11(2): 335-365.
- Anscombe, G.E.M. (1957). *Intention*. Cambridge, Massachusetts: Harvard University Press.
- Arcangeli, M. & Dokic, J. (2018). Affective Memory: A Little Help From Our Imagination. In Michaelian, K., Debus, D. & Perrin, D. (eds.) *New Directions in the Philosophy of Memory*. Routledge.
- Bae, Gi-Yeul, et al. (2015). Why Some Colors Appear More Memorable Than Others: A Model Combining Categories and Particulars in Color Working Memory. *Journal of Experimental Psychology* 144 (4): 744-763.
- Bailey, O. (2022). Empathy and the Value of Humane Understanding. *Philosophy and Phenomenological Research* 104 (1): 50-65.
- (2023a). Empathy, Sensibility, and the Novelist's Imagination. In Engisch, P. & Langkau, J. (eds.) *The Philosophy of Fiction: Imagination and Cognition*. Routledge.
- (2023b). What Must Be Lost: On Retrospection, Authenticity, and Some Neglected Costs of Transformation. *Synthese* 201(6): 189.
- Balcerak Jackson, M. (2013). Reasoning as a Source of Justification. *Philosophical Studies* 164 (1): 113-126.
- Bernecker, S. (2010). *Memory: A Philosophical Study*. Oxford: Oxford University Press.
- Boyle, M. (2024). *Transparency and Reflection*. Oxford University Press.
- Brainerd, C.J., & Reyna, V.F. (2002). Fuzzy-Trace Theory and False Memory. *Current Directions in Psychological Science* 11: 164-169.
- Byrne, A. (2010). Recollection, Perception, Imagination. *Philosophical Studies* 148 (1): 15-26.
- Campbell, J. (2002). *Reference and Consciousness*. Oxford: Clarendon Press.
- Conway, M.A. (2005). Memory and the Self. *Journal of Memory and Language* 53 (4): 594-627.
- Coplan, A. (2011). Understanding Empathy: Its Features and Effects in Coplan, A. & Goldie, P. (eds.) *Empathy: Philosophical and Psychological Perspectives*. Oxford University Press.

- D'Arms, J. (2000). Empathy and Evaluative Inquiry. *Chicago-Kent Law Review* 74: 1467-1500.
- De Brigard, F. (2014a). Is Memory for Remembering? Recollection as a Form of Episodic Hypothetical Thinking. *Synthese* 191 (2): 155-185.
- (2014b). The Nature of Memory Traces. *Philosophy Compass* 9 (6): 402-414.
- (2017). Nostalgia and Mental Simulation, in Gottlieb, A. (ed.) *The Moral Psychology of Sadness*. Rowman & Littlefield: 155-181.
- De Brigard, F., Addis, D., Ford, J., Schacter, D., & Giovanello, K.S. (2013). Remembering What Could Have Happened: Neural Correlates of Episodic Counterfactual Thinking. *Neuropsychologia* 51 (12): 2401-2414.
- Debus, D. (2007a). Being Emotional about the Past: On the Nature and Role of Past-Directed Emotions. *Noûs* 41 (4): 758-779.
- (2007b). Perspectives on the Past: A Study of the Spatial Perspectival Characteristics of Recollective Memories. *Mind & Language* 22 (2): 173-206.
- (2008). Experiencing the Past: A Relational Account of Recollective Memory. *Dialectica* 62 (4): 405-432.
- (2014). 'Mental Time Travel': Remembering the Past, Imagining the Future, and the Particularity of Events. *Review of Philosophy and Psychology* 5 (3): 333-350.
- Donnellan, K.S. (1966). Reference and Definite Descriptions. *Philosophical Review* 75 (3): 281-304.
- Dudai, Y. (2012). The Restless Engram: Consolidations Never End. *Annual Review of Neuroscience* 35 (1): 227-247.
- Fernandez, J. (2016). Epistemic Generation in Memory. *Philosophy and Phenomenological Research* 92 (3): 620-644.
- Frise, M. (2018). Forgetting. In Michaelian, K., D. Debus, & D. Perrin (Eds.), *New Directions in the Philosophy of Memory*. New York, NY: Routledge.
- Goldie, P. (2005). Imagination and the Distorting Power of Emotion. *Journal of Consciousness Studies* 12 (8-10): 127-139.
- (2011). Anti-Empathy. In Coplan, A. & Goldie, P. (eds.) *Empathy: Philosophical and Psychological Perspectives*. Oxford University Press.
- Hoerl, C. (2014). Remembering Events and Remembering Looks. *Review of Philosophy and Psychology* 5 (3): 351-372.
- Hopkins, R. (2014). Episodic Memory as Representing the Past to Oneself. *Review of Philosophy and*

- Psychology* 5 (3): 313–331.
- (2018). Imagining the Past. In Macpherson, F. and Dorsch, F. (eds.) *Perceptual Imagination and Perceptual Memory*. Oxford University Press.
- Kind, A. (2020). What Imagination Teaches. In Lambert, E. & Schwenkler J. (eds.) *Becoming Someone New: Essays on Transformative Experience, Choice, and Change*. Oxford University Press.
- (2021). Bridging the Divide: Imagining Across Experiential Perspectives. In Badura, C. & Kind, A (eds.) *The Epistemic Role of Imagination*. Routledge.
- Korsgaard, C.M. (1996). *The Sources of Normativity*. Cambridge University Press.
- Kung, P. (2010.) Imagining as a Guide to Possibility. *Philosophy and Phenomenological Research* 81 (3): 620-663.
- Kvart, I. (1989). Divided Reference. *Midwest Studies in Philosophy* XIV: 140-179.
- Langkau, J. (2021). On Imagining Being Someone Else. In Badura, C. & Kind, A (eds.) *The Epistemic Role of Imagination*. Routledge.
- Langland-Hassan, P. (2016). On Choosing What to Imagine. In Kind, Amy and Kung, Peter (eds.) *Knowledge Through Imagination*. Oxford University Press.
- (2021). What Sort of Imagining Might Remembering Be? *Journal of the American Philosophical Association* 7 (2): 231-251.
- (2022). Propping Up the Causal Theory. *Synthese* 200 (2): 1-27.
- Lackey, J. (2005). Memory as a Generative Epistemic Source. *Philosophy and Phenomenological Research* 70 (3): 636-658.
- Leslie, S.J. (2012). Generics. In Russell, G. & Fara, D (eds.), *Routledge Handbook of Philosophy of Language*. London: Routledge.
- Lewis, David. (1990). What Experience Teaches. In Lycan, W.G. (ed.), *Mind and Cognition*. Cambridge, Massachusetts: Basil Blackwell.
- Loftus, E. (1996). *Eyewitness Testimony*. 2nd ed. Cambridge, MA: Harvard University Press.
- (2005). Planting Misinformation in the Human Mind: a 30-year Investigation of the Malleability of Memory. *Learning & Memory* 12 (4): 361-366.
- Martin, C.B., & Deutscher, M. (1966). Remembering. *Philosophical Review* 75 (2): 161-196.
- Martin, M.G.F. (2001). Out of the Past: Episodic Recall as Retained Acquaintance. In Hoerl, C., & T. McCormack (eds.), *Time and Memory*. Oxford: Clarendon Press.
- (2002). The Transparency of Experience. *Mind and Language* 17: 376-425.

- McCarroll, C. (2018). *Remembering from the Outside: Personal Memory and the Perspectival Mind*. New York, NY: Oxford University Press.
- (2020). Remembering the Personal Past: Beyond the Boundaries of Imagination. *Frontiers in Psychology* 11: 585352.
- McCarroll, C., & Sutton, J. (2017). Memory and Perspective. In Bernecker, S., & K. Michaelian (eds.), *The Routledge Handbook of Philosophy of Memory*. New York, NY: Routledge.
- McCormack, T. (2001). Attributing Episodic Memory to Animals and Children. In Hoerl, C., & T. McCormack (eds.), *Time and Memory*. Oxford: Clarendon Press.
- McGrath, M. (2017). Knowing What Things Look Like. *Philosophical Review* 126 (1), 1-41.
- (2007). Memory and Epistemic Conservatism. *Synthese* 157(1), 1-24.
- Michaelian, K. (2011). The Epistemology of Forgetting. *Erkenntnis* 74 (3): 399-424.
- (2011). Generative Memory. *Philosophical Psychology*, Vol. 24 (3), pp. 323-342.
- (2016). *Mental Time Travel: Episodic Memory and our Knowledge of the Personal Past*. Cambridge, Massachusetts: The MIT Press.
- (2023). If Remembering Is Imagining, Then What Is Forgetting? In Anja B.A. & Ferran, I.V. (eds.) *Philosophical Perspectives on Memory and Imagination*. Routledge.
- Moran, Richard. (2012). *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press.
- Mori, N. (2009). The Schema Approach: A Dynamic View on Remembering. In *Dynamic Process Methodology in the Social and Developmental Sciences*. New York: Springer.
- Openshaw, J. (2022). Remembering Objects. *Philosophers' Imprint* 22 (11): 1-20.
- Openshaw, J., and Michaelian, K. (Forthcoming.) Reference in Remembering: Towards a Simulationist Account. *Synthese*.
- Papineau, D. (2006). Phenomenal and Perceptual Concepts. In Alter, T., & S. Walter (eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press.
- Paul, L.A. (2014). *Transformative Experience*. Oxford University Press.
- (2017). First Personal Modes of Presentation and the Structure of Empathy. *Inquiry* 60 (3): 189-207.
- Peacocke, C. (1985). Imagination, Conceivability and Possibility. In Foster, J., & H. Robinson (eds.),

- Essays on Berkeley: A Tercentennial Celebration*. Oxford: Clarendon Press.
- Perrin, D. (2019). Le Contenu du Souvenir Épisodique: Une Singularité non Fondée sur l'Accointance. *Les Études Philosophiques* 193 (3): 479-496.
- Recanati, F. (2007). *Perspectival Thought: A Plea for (Moderate) Relativism*. Oxford: Oxford University Press.
- Robin, J., and Moscovitch, M. (2017). Details, Gist, and Schema: Hippocampal-Neocortical Interactions Underlying Recent and Remote Episodic and Spatial Memory. *Current Opinion in Behavioral Sciences* 17: 114-123.
- Robins, S.K. (2016a.) Misremembering. *Philosophical Psychology* 29 (3): 432-447.
- (2016b). Representing the Past: Memory Traces and the Causal Theory of Memory. *Philosophical Studies* 173 (11): 2993–3013.
- (2020). Defending Discontinuism, Naturally. *Review of Philosophy and Psychology* 11: 469-468.
- Roediger, H., and McDermott, K. (1995). Creating False Memories: Remembering Words That Were Not Presented in Lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21: 803-814.
- Sartre, J.P. (1940; 2004). *The Imaginary: A Phenomenological Psychology of the Imagination*. Tr. Webber, Jonathan. London and New York: Routledge.
- Schacter, D. (2021). *The Seven Sins of Memory: How the Mind Forgets and Remembers*. Boston: Mariner Books.
- Schacter, D., and Addis, D. (2007). On the Constructive Episodic Simulation of Past and Future Events. *Behavioral & Brain Sciences* 30: 299-351.
- Schurigin, M.W. (2008). Visual Memory, the Long and Short of it: A Review of Visual Working Memory and Long-Term Memory. *Attention, Perception, and Psychophysics* 80 (5): 1035-1056.
- Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Siegel, S., and Silins, N. (2017). The Structure of Episodic Memory: Ganeri's 'Mental Time Travel and Attention'. *Australasian Philosophical Review* 1(4): 374-394.
- Soteriou, M. (2013). *The Mind's Construction: The Ontology of Mind and Mental Action*. Oxford: Oxford University Press.
- (2018). The Past Made Present: Mental Time Travel in Episodic Recollection. In Kourken M., Dorothea D. & Denis P. (eds.) *New Directions of Research in the Philosophy of Memory*. New York: Routledge.
- Squire, L.R., Stark, C.E.L., & Clark, R.E. (2004). The Medial Temporal Lobe. *Annual Review of Neuroscience* 27 (1): 279-306.

- Stanley, J., & Williamson, T. (2001). Knowing How. *Journal of Philosophy* 98: 41-44.
- Sutton, J. (1998). *Philosophy and Memory Traces: Descartes to Connectionism*. Cambridge University Press.
- Teroni, F. (2017). The Phenomenology of Memory. In Bernecker, S., & K. Michaelian (eds.), *The Routledge Handbook of Philosophy of Memory*. New York, NY: Routledge.
- Tulving, E. (1972.) Episodic and Semantic Memory. In Tulving, E. and Donaldson, W. (eds.) *Organization of Memory*. New York: Academic Press.
- Velleman, D.J. (1996). Self to Self. *The Philosophical Review* 105(1): 39-76.
- Vendler, Z. (1984). *Matter of Minds*. Oxford: Oxford University Press.
- Williamson, T. (1994). *Vagueness*. London; New York: Routledge.
- Williamson, T. (2000). *Knowledge and Its Limits*. Oxford: Clarendon Press.
- Williamson, T. (2007). *The Philosophy of Philosophy*. Malden, MA: Blackwell Publishing.
- Williamson, T. (2016). Knowing By Imagining. In Kind, A. and Kung, P. (eds.) *Knowledge Through Imagination*. Oxford: Oxford University Press.
- Wittgenstein, L. (1958). *Blue and Brown Books*. Harper & Row, Publishers. New York, NY.
- Zhou, C., et al. (2022). Categorical Bias as a Crucial Parameter in Visual Working Memory: The Effect of Memory Load and Retention Interval. *Cortex* 154: 311-321.