

# Imagining the Actual

Daniel Munro

*University of Toronto*

© 2021 Daniel Munro

*This work is licensed under a Creative Commons  
Attribution-NonCommercial-NoDerivatives 3.0 License.  
<[www.philosophersimprint.org/021017/](http://www.philosophersimprint.org/021017/)>*

## 1. Introduction

Far from being relegated to representing only the non-actual, the imagination is a cognitive faculty that can be used to put us in touch with reality. It does so when we engage in the cognitive capacity I call *actuality-oriented imagining*. I'll begin with two illustrative examples. First, suppose you're asked how many windows are on the outside of your house, a question you've never explicitly entertained before. The answer doesn't immediately come to mind, so you pause momentarily before responding. During this pause, you visualize the outside of your house; you mentally rotate the house in your visual image while counting the number of windows on each side. You then respond, "It has ten windows."

Now suppose you're asked whether your balding uncle has more or less than half his head of hair remaining, again a question you've never explicitly entertained before. Your uncle isn't currently anywhere in the immediate vicinity, so you pause to visualize his head before answering. You mentally examine his head from various angles to take note of how much hair your image contains, estimating that he has less than half of his hair remaining. You then respond to the question accordingly.

Assuming you've had enough prior perceptual experiences of your house and of your uncle, this use of mental imagery seems like the natural way to go about answering these questions. While we often associate the imagination with our capacity to represent hypothetical or fictional states of affairs, cases like these make salient that we also use it to represent things as they are in the actual world. Such cases are my focus in this paper. More specifically, I'm focused on cases in which, as in the above two examples, a subject intentionally imagines as a means of coming up with an answer to some currently salient question about the actual world. In such cases, one doesn't form a mental image that merely happens to be of, or resemble, something actual; rather, one uses one's imagination in a way that's aimed at representing the actual.

This paper considers two questions about actuality-oriented imagining. The first, which I'll call the "content determination" question, asks how to specify the factors that determine which object(s) an actuality-oriented mental image represents. In other words, when I imagine some actual object (e.g., my house), what makes it the case that my mental image represents *that* object, as opposed to some distinct one (e.g., a fake house façade)? The second, the "success conditions" question, asks about the success conditions of actuality-oriented imaginings — under what conditions are such imaginative states of mind successful? My goal in this paper is to give partial answers to these questions and, in doing so, to reveal ways that actuality-oriented imagining is interestingly distinct from other, similar cognitive capacities.

In §2, I'll first motivate the claim that we need to answer these questions in a way that's specific to actuality-oriented imagining rather than borrowing existing answers about similar capacities. Specifically, I'll argue that although actuality-oriented imagining has similarities to both perception and other kinds of imagining, the kinds of answers typically given for those faculties seem intuitively not to be applicable to actuality-oriented imagining. Section 3 then more concretely probes the cognitive structure of actuality-oriented imagining, by situating it within the "predictive processing" framework from cognitive science. This framework helps to bring out the fact that actuality-oriented imagining has both perceptual and active elements, since it's generated using perceptual cognitive mechanisms but is also a particular kind of mental action. After elucidating these two aspects of actuality-oriented imagining, §4 will leverage this account to help fill out what makes actuality-oriented imagining distinctive, in terms of content determination and success conditions.

The ultimate aim of this paper is to show that actuality-oriented imagining is a cognitive capacity that is philosophically interesting in its own right, warranting individualized investigation. In this spirit, §5 concludes by sketching two implications, for the philosophy of imagination more generally, of this paper's arguments.

## 2. Existing answers about similar capacities

This section considers the plausibility of adopting answers, to both the content determination and success conditions questions, that are often given to analogous questions about similar cognitive capacities. One such capacity I'll consider is perception. It's common for philosophers to compare imagination with perception, given their similarity of phenomenal character and underlying cognitive mechanisms.<sup>1</sup> Actuality-oriented imagining, in particular, has an affinity with perception that other kinds of imagining don't, given that it's directed at representing the actual world as it is in the present. In addition to perception, I'll consider answers to each question that are adopted from common views about imagination more generally. Section 2.1 will consider answers to the content determination question from both perception and other kinds of imagination; §2.2 will consider answers to the success conditions question from the same.

Each of the next two subsections starts by describing a case of attempted actuality-oriented imagining in which it seems intuitively that a subject *mis*-imagines, in different ways. The case in §2.1 involves a kind of imaginative failure that we should expect a correct answer to the content determination question to explain; the same goes for the case of imaginative failure in §2.2 and the success conditions question. I'll argue that existing answers to each question, both about perception and imagination, fail to explain what goes wrong in these cases. I'll also gesture towards what these answers intuitively seem to leave out, that is, what intuitively seems to be distinctive of actuality-oriented imagining that generates these cases of imaginative failure. In §4, I'll ultimately return to these intuitions and aim to both vindicate them and make them more precise.

1. Philosophers sometimes distinguish *sensory* imagining, which involves mental imagery, from purely *propositional* imagining, which is something like an imaginative analogue of belief. It's controversial whether there's such a thing as purely propositional imagining (cf. Kind 2001; Balcerak Jackson 2016), but I'm not taking a stance on this issue. I'm concerned in this paper only with imagining that involves imagery, so I leave open whether there are also non-sensory imaginings.

*2.1. On content determination*

Here's the first case of attempted actuality-oriented imagining in which a subject seems to mis-imagine:

**WINDOWS:** Peggy is a realtor with many houses currently on the market. One of these, a house called Alpha Manor, is a charming brick cottage in the countryside. Another, called Beta Estate, is a palatial seaside mansion. She has seen both of these houses in person and in photos many times so is well-acquainted with each. One day, she receives a call from a potential buyer for Alpha Manor. This caller is particularly concerned with whether the interior of the house receives a lot of natural light so asks how many windows it has. Not having explicitly counted the number of windows before, and without any photos of the house nearby, Peggy pauses to form a visual image of Alpha Manor in order to count its windows. However, when she does so, she accidentally forms the image by drawing on her mental store of information about Beta Estate. The house in her mental image thus resembles Beta Estate and not Alpha Manor, though she takes her image to be of Alpha Manor. She counts the windows in her mental image and reports to the caller accordingly.

Intuitively, it seems here that Peggy fails to imagine the object she intends to imagine: she tries, but fails, to imagine Alpha Manor. It's not immediately obvious what object her imagining represents – if it represents Beta Estate while she mistakenly believes it represents Alpha Manor or if it fails to represent any particular house. Regardless, what's important here is the intuition that Peggy tries to imagine a particular, actual house but fails to do so. The question I'm concerned with in this subsection is: What is the necessary condition for content determination that Peggy fails to meet, such that she fails to represent Alpha Manor?

It's relatively easy to see why we can't just borrow the content determination conditions of perceptual experience to answer this question. Suppose I stand out on my street looking at my house. What makes it the case that my visual experience represents *this* house rather than some other object? While the exact details of how to answer this question are controversial, it seems that they must somehow be grounded in the fact of my visual system's occurrent perceptual relation to my house, a relation that requires an occurrent causal connection to an object in my perceptible environment. When I fail to represent some object in my environment, it's in virtue of the fact that I'm not perceptually "hooked up" to it in the right way. Unlike perception, actuality-oriented imagining clearly doesn't require this kind of occurrent perceptual connection, since Peggy can imagine Alpha Manor even when it's nowhere in her immediate vicinity. So, we can't appeal to the fact that Peggy doesn't stand in an occurrent perceptual relation to Alpha Manor to explain why she fails to imagine it.

Might we instead adopt an answer from existing literature on the content determination of imaginings? I'll now consider and reject what may be the most prominent such answer. One of the major contrasts between perception and imagination is that what we imagine is under our voluntary control to a degree that what we perceive isn't.<sup>2</sup> It might seem that this degree of control has implications for determining what a mental image represents – namely, because the content we imagine is up to us, what our imagery represents depends on what we intend to imagine and/or believe we are imagining. In this vein, views that I'll call "subjectivist" hold that our intentions about what to imagine and/or our beliefs about what we are imagining are sufficient for determining the object(s) our mental imagery represents. This sort of view is

2. This point has historically been emphasized by various philosophers, including Sartre and Wittgenstein (Balcerak Jackson 2018). There are, of course, ways in which we can control what we perceive (e.g., by controlling attention or where we direct our gaze), and we sometimes imagine things without explicitly choosing to (e.g., when our minds wander or we get some image stuck in our heads). Nevertheless, it seems intuitive that the content of our mental imagery is subject to our intentions about what to imagine in a way that perception isn't.

quite common in literature on the imagination, though it's stated in slightly different ways by different theorists (see, e.g., McGinn 2004, sec. 1.7; Dorsch 2012, ch. 3; Kung 2016, 626; Kind 2019, 166–167). I'll focus on it as the very general claim that one's beliefs and/or intentions about what one is imagining are sufficient for determining what one is imagining.

Reflection on examples initially shows why subjectivism is appealing. Consider the following passage from Paul Tidman (1994), who borrows an example from Ludwig Wittgenstein:

[S]uppose someone claims to imagine King's College on fire. It would be absurd to ask, "Are you sure it's King's College? Maybe you are just imagining a building that looks like King's College." Whether one is imagining King's College depends not on the image, but on what we take the image to be an image of. (301)

Tidman thus argues that what an image represents depends on what the imaginer takes it to represent. In general, the fact that two phenomenally identical images can represent different objects depending on what the imaginer takes them to represent suggests that there's a close link between an image's content and the imaginer's intentions and beliefs. A phenomenally identical mental image could represent either King's College or a fictional building that looks like King's College, a real house or a realistic fake house façade, and so on. Which object it really represents seems to depend on what I'm trying to imagine or believe myself to be imagining.

My aim here isn't to contest the subjectivist treatment for cases of non-actuality-oriented imagining, or even to argue that intentions and beliefs don't make any difference in actuality-oriented cases. However, WINDOWS shows that subjectivist conditions seem at least *insufficient* for determining what a mental image represents in actuality-oriented cases. Subjectivism about actuality-oriented imagining would straightforwardly imply that Peggy does imagine Alpha Manor, because she intends to imagine Alpha Manor and believes she is doing so. This

seems intuitively like the wrong description of WINDOWS. Subjectivism makes it very difficult for one to be wrong about what one is imagining, but this seems to be what happens to Peggy when she thinks she's imagining Alpha Manor. Against the above quote from Tidman, it seems perfectly sensible, at least in an actuality-oriented case like WINDOWS, to question whether Peggy really is imagining the object she thinks she's imagining.

Here's my first gesture towards the condition for the content determination of actuality-oriented imaginings that Peggy fails to meet: in actuality-oriented imagining cases, the stored information a subject draws on to form a mental image (at least partly) determines what that mental image represents. In other words, to imagine some object, it's necessary to draw on information about what *that* object is like rather than some distinct object. Peggy possesses information, acquired from past experiences, about both Alpha Manor and Beta Estate. However, when she tries to retrieve stored information about Alpha Manor in order to imagine it, she accidentally retrieves information about Beta Estate instead. That seems to be where she goes wrong in WINDOWS, and what makes it the case that she fails to imagine Alpha Manor.<sup>3</sup>

This apparent need to possess stored information about some object in order to represent it seems to distinguish actuality-oriented imagining from perception and from other kinds of imagining. Because perceptually representing some object involves coming into a perceptual causal relation to it, we don't necessarily need to possess any prior information about an object to perceive it. Instead, we can perceive novel objects by entering into that perceptual relation with them.<sup>4</sup> And, when in a context of non-actuality-oriented imagining, it

3. This suggests that there's a close affinity between actuality-oriented imagining and memory. I explore this in more detail in §3.

4. Readers already familiar with the predictive processing framework I'll subscribe to below in §3 might object here that, according to that framework, constructing perceptual representations involves drawing on stored information we already possess about the world. However, while this framework may indeed require that I already have some general stored information about the world in order to perceptually represent some house, it doesn't require that

seems we can arbitrarily stipulate that our imaginings represent particular objects even if we don't possess any individuating information about their properties. I may not have a clue what Socrates looked like, but that doesn't stop me from imagining some arbitrary man and just stipulating that I'm imagining Socrates. I can also use imagination for purposes like designing a new fictional creature that I have never conceived of before, such as by mentally combining body parts of various kinds of animals in a new way. When doing so, I can form a mental image of this new creature and then, subsequently, name it, even if I had no previous idea what it would end up looking like. It may be that all sensory imaginings are constructed by drawing on stored information of some kind — perhaps information about men in general in the Socrates example and information about various kinds of animals in the fictional creature example. Nevertheless, such examples still differ from WINDOWS, in which Peggy must draw on information about Alpha Manor in order to imaginatively represent it.<sup>5</sup>

Again, this has been only a gesture towards what's going wrong in WINDOWS, and part of my goal in §4 will be to precisify it.

### 2.2. *On success conditions*

Now I turn to the success conditions question. I'll start by considering another case of apparent mis-imagining, but, in this example, a subject seems intuitively not to meet actuality-oriented imagining's success conditions:

---

I already have information about that *particular* house for me to see it for the first time. I say more about these matters in §4.

5. In this subsection, I haven't totally ruled out that there exist non-actuality-oriented kinds of imagining to which this point about content determination also applies. There is such a huge range of types of imaginative acts — in modal epistemology, mindreading, fantasy, engagement with fiction, etc. — that I can't rule this out here. However, even if it's not *unique* to actuality-oriented imagining that representing an object requires drawing on stored information about it, this is at least a characteristic that sets it apart from many cases of imagining and from perception.

**DRAPES:** Ed purchases Alpha Manor and sets about doing some refurbishing on his new home. While visiting a home furnishing store, Ed decides on a whim to buy matching sets of new drapes for all of the house's windows. He wonders to himself how many sets of drapes he needs to buy. Since he hasn't counted Alpha Manor's windows before and doesn't have a photo with him, he pauses to try to form a mental image of Alpha Manor in order to count its windows. The house in his mental image resembles Alpha Manor closely, except that he mistakenly imagines it with an extra window — with eleven windows instead of its actual ten. He concludes that Alpha Manor has eleven windows so buys one too many sets of drapes.

Ed's mis-imagining seems to be of a different kind than Peggy's in WINDOWS. In that case, Peggy failed to imagine the object she intended to imagine. In DRAPES, it seems that Ed does imagine Alpha Manor but nevertheless mis-imagines in virtue of getting the number of windows on Alpha Manor wrong. So, we can ask: What condition(s), for an actuality-oriented imagining to successfully represent an object, does Ed fail to meet?

Initially, the most obvious candidate is that Ed's imagining is unsuccessful in virtue of failing to *veridically* represent Alpha Manor, given that he mis-imagines its number of windows. This would involve bringing the success conditions for actuality-oriented imagining in line with those for perception, in which we typically take a successful representation to be one that veridically represents its objects. This might seem especially plausible given that actuality-oriented imagining is like perception in representing things in the actual world.

The problem with this proposal is that full veridicality seems intuitively to be too demanding a standard. Consider a modified version of DRAPES in which Ed *does* imagine Alpha Manor with the right number of windows and in which his mental image fairly closely resembles Alpha Manor overall. But suppose that he nevertheless gets various

other details wrong — say, the windows are the wrong shape and size, the doorknob is on the wrong side of the front door, and the color of the house is too dark. Despite these inaccuracies, it doesn't seem like his imagining is therefore unsuccessful. Rather, it seems like a general feature of actuality-oriented imagination that we usually, if not always, imagine with various inaccuracies of this kind. So, it doesn't seem right to apply perception's success conditions to actuality-oriented imagining.

Is there an existing alternative proposal about the success conditions of imagination that's more promising? I think the most obvious place to look is existing literature on the *function* of imagination — specifically, to literature about the function of imagination in cases where it's being used for epistemic purposes (rather than for merely fantasizing or daydreaming, e.g.). This is because we can determine a cognitive faculty's success conditions by first determining its function. If, for example, we take perception's cognitive function to involve tracking things as they actually are in the world around us, we can think of a given instance of perception as successful when it represents veridically (cf. Graham 2014). Similarly, recent work on how to distinguish successful cases of episodic remembering from cases of mis-remembering has focused on first understanding episodic memory's function, then arguing that mis-rememberings are those that fail to fulfill this function (cf. De Brigard 2014; Michaelian 2016b). In this spirit, it seems promising to look to existing philosophical views about the function of imagination to help answer the success conditions question.

Unfortunately, though, prominent existing views about the function of imagination don't seem applicable to cases of actuality-oriented imagining. It's common to contrast imagining with perception by claiming imagination's function is to cognize possibilities rather than actuality. This tradition is grounded in influential work by, for example, Stephen Yablo (1993), who takes one of the primary epistemic values of the imagination to be delivering knowledge about what's possible. In this vein, Timothy Williamson (2016) likens imagination to "attention to possibilities." Similarly, Magdalena Balcerak Jackson (2018)

argues that the function of imagination is to recreate possible perceptual experiences, which grounds its ability to tell us about what's metaphysically possible. If the function of imagination is to represent possibilities, then an imagining is successful when it represents a possible state of affairs. While this may indeed be true of many cases of imagining, it doesn't intuitively capture the success conditions for an actuality-oriented case like DRAPES. That's because Ed's mental image *does* reflect a possible way Alpha Manor could be: there are fairly nearby possible worlds in which Alpha Manor was built with eleven windows instead of ten, as well as worlds in which Ed does some home renovations to add an eleventh window. Nevertheless, his imagining seems unsuccessful. Although full veridicality seemed like too strict a standard, it seems like mere possibility is not strict enough — it's relevant to the success of Ed's imagining whether he imagines the house's actual number of windows, even if he gets other details wrong.

So, actuality-oriented imagining fits comfortably neither within perception's success conditions nor within a prominent existing framework for thinking about the function of imagination. Rather, it seems that we instead need somehow to relativize our success conditions for Ed's imagining to the purpose or goal behind it. In other words, it seems that, to be successful, an actuality-oriented imagining need only veridically represent an object *in respects relevant to the question(s) a subject is aiming to answer* about that object. In DRAPES, Ed is wondering about the number of windows on Alpha Manor, so it's relevant that he gets this property correct. This is a stricter success condition than for imaginings aimed at representing what's merely possible, such as those used to investigate metaphysical possibility. It's also a less strict success condition than for perception, for which we expect a higher standard of veridicality.

Much as in §2.1, that's a somewhat vague, intuitive characterization of what makes actuality-oriented imagining distinctive, relative to the success conditions question. Section 4 will aim to flesh this out more fully. First, though, §3 will elucidate the cognitive structure of actuality-oriented imagining in more detail.

### 3. Actuality-oriented imagining: Perceptual and active

“Predictive processing” (hereafter cited as PP) is one of the dominant frameworks for studying perception and action in recent cognitive science and philosophy of mind (for thorough overviews, see Hohwy 2013; Clark 2016). In this section, I’ll draw on elements from the PP framework to elucidate the structure of actuality-oriented imagining cases.<sup>6</sup> Section 3.1 first explains perception within PP, then extends these perceptual mechanisms to the construction of mental imagery. Section 3.2 then explains some key elements of the framework’s approach to action, especially certain kinds of mental action. Finally, §3.3 applies both perceptual and active aspects of PP to analyze cases of actuality-oriented imagining. First, though, a few general, preliminary points about the PP framework.

The rest of this section will appeal to two key notions at the core of PP. The first is the brain’s “generative model,” which constitutes its store of information about the world. This is usually framed in Bayesian terms, as a distribution of prior probabilities over hypotheses about states of the world. These include hypotheses of all sorts: where various objects and events in the world are located; what properties these objects have; the kinds of causal laws and regularities things in the world are, or tend to be, governed by; and so on. This model thus represents both the states of various parts of the world and information relevant for predicting how those states will evolve over time. It thereby structures our perceptual systems’ expectations about what we’ll encounter as we observe and interact with the world, by constantly generating predictions about what the perceived world will be like. This model is then constantly revised and updated as we perceive

6. Although the PP framework is—unsurprisingly—not uncontroversial, I won’t take up much space reviewing evidence and arguments in favor of it. I take it that the framework is popular and mainstream enough in the relevant sciences to justify using it as a starting point. I’m also appealing mainly to elements of the framework that are shared among various proponents of PP rather than to any particular, idiosyncratic developments of the framework by specific theorists.

the world and learn more about it.<sup>7</sup> In what follows, I will write as if the generative model’s hypotheses about the world are sub-personal states. It will thus sound as if states traditionally thought of as “cognitive”—in particular, person-level beliefs—are distinct from the cognitive architecture described by PP. However, I’m primarily doing this for simplicity. Many of PP’s strongest proponents (including Hohwy 2013; Clark 2016) extend the framework to both personal and sub-personal processes. Even if they’re right about this, though, I think it’s plausible to nevertheless maintain that person-level states are importantly distinct from sub-personal hypotheses, in a way that (at least roughly) tracks the more traditional distinction (for discussion, see Dewhurst 2017; Drayson 2017; Macpherson 2017).

The second core notion is what PP sees as the brain’s main organizing principle: “prediction error minimization.” This essentially means that the brain is constantly trying to correct errors in its predictions about the world, that is, to minimize discrepancies between the way the world is and the way its generative model says the world is. A more accurate model of the world, and of the way the states of the world will evolve over time, will result in more accurate predictions about what one will encounter as one goes about navigating the world. A less accurate model will result in discrepancies between what the brain expects and what it encounters. Intuitively, it makes sense that prediction error minimization would be (at least one of) the brain’s main evolved function(s), given that it’s conducive to an organism’s ability to satisfy basic needs, such as survival, finding nourishment, and reproducing. The more accurate one’s model and predictions, the

7. It’s also important for fully understanding the framework that this model is organized *hierarchically*: upper layers of this hierarchy contain hypotheses about more abstract properties over longer timescales (e.g., about ordinary objects and causal interactions between them), while lower layers contain less abstract properties over shorter timescales (e.g., about properties like colors, edges, and motion at short timescales). When the brain makes predictions about its environment, hypotheses at each layer of the hierarchy act as priors for the layer below. My arguments don’t depend directly on these facts about hierarchical organization, so I’ll gloss over this part of the framework in what follows.

less likely one is to encounter surprising, potentially dangerous or life-threatening situations for which one is unprepared. Furthermore, a more accurate model of the world means more accurate predictions about matters such as where to find food, potential mates, and so on. And one's ability to effectively act in the world, in ways conducive to one's self-interest, is greatly enhanced by an ability to accurately predict the consequences of one's actions.

### 3.1. *Perception and imagery*

In what follows, I'll mostly focus on vision, though the account is also meant to be adaptable to other perceptual modalities.

It's perhaps easiest to understand how PP views perception by contrasting it with more "traditional" scientific accounts of perception. On traditional accounts of vision, for example, perceptual experience is the result of the brain detecting information about the world that it receives in retinal input. We receive retinal stimulation from the external environment, from which the brain extracts information and assembles it into coherent perceptual representations. This is a "bottom-up" process because it involves passing information up from the world to the brain.

In contrast, PP views visual experience as largely a product of "top-down" processing. Instead of the more passive process of receiving input and detecting features of the environment it conveys, perceptual processing is a more active, constructive process, in which the brain uses the information in its generative model to construct representations of one's environment. As I navigate the world, my brain generates predictions about what I will encounter, based on the generative model's set of most likely hypotheses about what's in my perceptible environment. As this occurs, I still receive sensory data from the world in a bottom-up way, but the role of this data is to function as evidence against which the brain tests its top-down hypotheses. If the brain's predictions are the hypotheses that best explain incoming sensory input, they're deemed successful. If sensory evidence conflicts with the brain's predictions, error signals are generated to indicate that these

predictions must be revised; the brain then tries to revise its hypotheses to eliminate prediction error. So, visual processing is a process of the brain trying to minimize error in its perceptual predictions. A stable perceptual representation emerges when the brain has settled on the set of hypotheses that best minimizes prediction error. This framework thus differs from more traditional accounts of perception in that the construction of perceptual representations is driven by top-down predictions, although perceptual processing still involves the interaction of both top-down and bottom-up elements.

As this process occurs, the brain's generative model is also being updated, since the process of settling on the set of most likely hypotheses is also a process of revising the brain's model of the world. The revised model then becomes the basis of subsequent predictions. The generative model is thus constantly collecting more and more information that's used to inform future prediction and error minimization. When one sees some part of the world one has already perceived before, then, the brain's initial predictions are drawn from information it already has about that part of the world — that is, perceptual processing would begin by drawing on information the generative model already has about the states of that part of the world, after which its predictions are confirmed against sensory evidence and revised as needed.

That's a basic sketch of perception within PP. Though it leaves much out, it's sufficient for my purposes.<sup>8</sup> I'll turn now to mental imagery.

Some of PP's proponents, including Andy Clark (2016) and Michael D. Kirchhoff (2018), have argued that the mechanisms involved in perception can be extended to explain the construction of mental imagery. It's already widely thought that there is significant overlap between the cognitive mechanisms that produce perceptual experiences and mental imagery (cf. Clark 2016, ch. 3). Given that PP

8. One important piece I left out is the role that "precision weighting" of bottom-up signals and top-down hypotheses plays in this picture: in particular, how this is thought to relate to perceptual *attention*. Without these pieces, the framework is significantly oversimplified — see Clark (2016, ch. 2) for much detailed discussion.



describes perceptual experience as constructed top-down, and given that mental imagery is also generated in a top-down, endogenous way, PP seems well-positioned to be extended to an account of imagery construction.

In PP, vision involves the brain using its store of existing information to generate perceptual predictions. These hypotheses are then confirmed against or revised on the basis of sensory input. But if the same kind of top-down image-construction process could also be implemented in a way that's shielded from correction on the basis of bottom-up sensory evidence, we would have mechanisms that seem capable of constructing mental imagery. In other words, PP already says that the brain is capable of using its store of information about the world to endogenously generate perceptual representations. If that's true, it's plausible that mental imagery is generated by drawing on the same store of information, just in a way that's, as Clark (2016) puts it, "insulated from" revision on the basis of sensory evidence. The result would be that the brain can generate imagery with a very similar phenomenology and representational format to perceptual experience — which sounds much like mental imagery. This is a process in which "imagining is essentially *reusing* some of the same prior probabilities that are generated, tuned and maintained by the agent when perceptually engaging with the world," bringing back to mind predictions that had previously been revised during perception (Kirchhoff 2018, 765).<sup>9</sup>

More specifically, this kind of account at least seems well-positioned to describe mental imagery *of the states of affairs one's generative model says are most likely*. In perceiving, my brain uses its model to

9. One important question this line of thought raises is how to explain the differences in phenomenology between perception and imagination, particularly perception's richness and phenomenal directness. Clark (2016, ch. 3) attempts to address this question. Although the connection between phenomenology and PP's descriptions of the brain are important and interesting, I'm using PP here primarily as a computational framework for modeling perception and imagery, not as a framework for describing phenomenology. So, I'll set aside these questions. For general discussion of the connection between PP and conscious phenomenology, see Hohwy (2013).

bring to mind the most likely hypotheses about the part of the world I'm currently looking at, which are then confirmed and revised against sensory evidence. If we can generate mental imagery using the same top-down mechanisms involved in perception (minus revision against bottom-up signals), this would also involve bringing to mind the most likely hypotheses about whatever part of the world one is imagining. In other words, when I try to imagine some part of the actual world, my brain constructs a representation of what it expects I *would* see were I actually perceiving that part of the world. When I'm imagining some part of the actual world that I've experienced before (and have no reason to think has changed), this would be a relatively simple process of bringing to mind information that was learned in past experience. Of course, we're also able to imagine false and fantastical scenarios. Accounting for all kinds of mental imagery would require extending the PP framework to explain how we can bring to mind contents that go far beyond what the generative model says is most likely. But since my topic in this paper is actuality-oriented imagining, this extension would go beyond my present needs.

I'll thus adopt this account of the construction of imagery representing parts of the actual world one has previously experienced (or, more precisely, of how my generative model takes the actual world to be). The key point here is that the construction of such imagery is far from a process of generating *new* information, as we might expect many uses of mental imagery to be. The imagination is often used more creatively, such as when we use it to combine our existing beliefs in novel ways (cf. Kind 2018) or to design novel fictional worlds. However, imagery of the actual world is formed just by bringing to mind the existing, rich sets of information we have collected during previous experiences, without transforming or altering that information. Such imagery is essentially a way of accessing the brain's existing model of some part of the world, even when we're not in current perceptual contact with that part.

I said above that PP sees the brain as fundamentally oriented towards prediction error minimization. So far, the account of how

mental imagery is constructed doesn't obviously imply anything about how imagining helps accomplish this — in fact, as it stands, it sounds more like imagining is inert in this sense, given that it's just a matter of bringing back to mind previously updated hypotheses without revising them in response to sensory evidence. In §3.3, I'll explain how actuality-oriented imagining is oriented towards prediction error minimization. First, though, I need to get some elements of PP's account of action on the table.

### 3.2. *Action: Physical and mental*

This subsection explains some components of the PP framework's account of action, though it once again glosses over many of the fundamentals and fine details. I focus just on the points relevant for giving an account of actuality-oriented imagining in §3.3. For more detailed general explanations of action in the PP framework, see Clark (2016, esp. ch. 4).

Suppose you've just moved to a neighborhood in walking distance from your office. You go to walk from your house to your office for the first time. As you do so, your brain's generative model also continuously predicts the actions you're performing, by predicting the states of your body in much the same way that it predicts other states of the world. Your brain would thus hypothesize that you're currently walking to your office, as well as predict the various sub-actions and bodily movements that are constitutive of this. As long as you continue to perform the actions your generative model says you're performing, there won't be any discrepancies between what it says about the states of your body and what the incoming flow of sensory evidence says about the states of your body (sensory evidence coming from, e.g., proprioceptive signals). If you were to suddenly stop performing the action (because, e.g., you tripped and fell on the street), the incoming sensory evidence would reflect this, and your generative model's hypotheses about what you're now doing would be revised. Thus, minimization

of perceptual prediction errors, regarding predictions about your own actions, is constantly occurring as you act.<sup>10</sup>

Now suppose that before you started your walk to the office, you consulted a map and tried to memorize the directions but brought the map along just in case you needed to refresh your memory. Eventually, you get to an intersection and find that you can't remember which way to turn next. In other words, although your current goal of walking to your office hasn't changed, the appropriate sub-actions constitutive of achieving this goal have become unclear. Furthermore, as you suddenly stop walking because you become uncertain about what to do next, a mis-match arises between the action your generative model says you're performing — walking to your office — and what current sensory input says you're doing — standing still on the street corner. Due to this discrepancy, your brain encounters prediction error.

The predictive brain is constantly attempting to minimize prediction error so must now take some step to eliminate it. One option is just to abandon the action of walking to your office while revising the prediction that you're doing so to bring it in line with the fact that you've stopped walking. However, another option is to perform what's called an "epistemic action" (Friston et al. 2015, 2016; Pezzulo and Nolfi

10. PP actually says that the links between action, predictions about one's actions, and prediction error minimization are much tighter than I've described here. PP doesn't merely say that my brain attempts to minimize errors in predictions about my own actions; it makes the stronger claim that actions *are* a kind of prediction error minimization. Prior to performing an action, my generative model predicts that I'm now performing it, with a set of hypotheses about my bodily movements. When initially generated, that set of hypotheses is false — it conflicts with current sensory input so results in prediction error. One way to eliminate this error is to revise the false hypotheses to reflect that I'm not currently acting (the kind of process involved in perception — revising the brain's hypotheses to accommodate sensory evidence). But another way is to *act* to make my brain's hypotheses true — that is, carry out the action my brain predicts I'm performing. This is known as "active inference" and is the core of PP's account of action. What I have said in this subsection is consistent with the full PP account, just weaker: I claimed that minimization of prediction error about my own actions occurs *as I act*, though I stopped short of saying that this is what action *is*. This is mainly because the full account introduces complexity that's unnecessary for my arguments in the rest of the paper.

2019). Rather than directly contributing to achieving the goal one is currently carrying out, the function of an epistemic action is to reduce uncertainty about *how* to achieve that goal. In the present example, the most obvious such action would be pulling out and consulting your map. An epistemic action allows one to, subsequently, continue acting in accordance with one's current hypotheses about how one is acting. It's thus another way to eliminate the prediction error arising from discrepancies between one's hypotheses and current bodily sensory evidence, one that doesn't require giving up one's current goals.

Pulling out and reading a map is a *physical* epistemic action — it involves using your body to act on the world to collect information. We can also conceptualize certain *mental* actions on a similar model (Pezzulo et al. 2016; Metzinger 2017; Pezzulo 2017). There's a very wide range of kinds of mental actions that could be included in this category; I'll be very general here, while the next subsection focuses more narrowly on how actuality-oriented imagining fits in. (For discussion of unifying various kinds of mental epistemic actions, see especially Metzinger 2017.) Consider a similar example of walking to your office and arriving at a point where you don't have the next set of directions memorized, but this time you've forgotten to bring your map along. You *could* perform a physical epistemic action, such as going back home to retrieve your map or asking someone on the street for directions. But another option is to engage in some kind of mental action, such as some kind of reasoning or recalling information. You might, for example, recall the fact that your office is to the southeast of your home and reason from there about which direction to go. Like physical epistemic actions, mental epistemic actions function to reduce uncertainty about what to do next, thereby allowing one to eliminate prediction error by continuing to act.

For my purposes in what follows, the key point in this subsection is this role that mental epistemic actions play in achieving the brain's core function of prediction error minimization. Discrepancies between what my brain predicts about my own actions and what I'm actually doing generate prediction error; this error can often be eliminated by

performing an epistemic action that reduces my uncertainty about how to act, and such actions are often mental.

The next subsection turns to applying features from this subsection and §3.1 to explain actuality-oriented imagining.

### 3.3. *Actuality-oriented imagining*

Actuality-oriented imagining has both perceptual and active characteristics. It has a perceptual phenomenal character and representational format and, like perception, represents actuality. At the same time, it's also a mental action that is undertaken intentionally and directed by a subject's current goals. We can describe the overall structure of actuality-oriented imagining cases by drawing on both perceptual and active aspects of the PP framework. I'll first do this in detail using the example of Peggy's imagining in WINDOWS, then extend the account to Ed's imagining in DRAPES.

First, we can apply the PP framework for action to describe the overall context in which an actuality-oriented imagining occurs, as well as the function it plays in that context. Take WINDOWS, in which Peggy is speaking to a potential buyer for Alpha Manor on the phone, answering her various questions. Peggy's generative model will predict her own actions accordingly: that she's speaking to the caller and answering her questions about Alpha Manor, along with the various actions that are constitutive of doing so (e.g., making particular vocalizations into the phone). Eventually, the caller asks Peggy how many windows are on Alpha Manor. As with every previous question, Peggy's generative model predicts that she answers this question.<sup>11</sup> However, doing so requires that Peggy have an explicit belief about the number of windows, so that she can communicate this information

11. Note that, generally, both parties to a conversation have an overarching expectation that questions will be answered appropriately, except in rare cases where, for example, a question has a false presupposition that the respondent must correct (cf. Stivers 2010). So, it makes sense to think that, by default, Peggy will expect herself to answer this rather straightforward question from the caller, unless it turns out that for some reason she's unable to do so and must therefore revise this initial expectation.

to the person on the phone. As per the case, though, Peggy has never formed an explicit belief about this.

Peggy therefore runs into a situation of uncertainty about how to act, so pauses. This pause generates prediction error, due to a discrepancy between the hypothesis that she's answering the caller's question and sensory evidence that she isn't doing so. One option to eliminate this prediction error is, of course, to revise the prediction that she's answering the caller's question and just not do so. What Peggy ends up doing instead is turning to an epistemic action to resolve her uncertainty. Peggy isn't in a position where she can easily perform a physical epistemic action, such as looking at Alpha Manor itself or pulling out a picture of it. She instead performs a mental epistemic action — namely, imagining Alpha Manor and counting the number of windows in it.

§3.1's PP model of mental imagery explains how this mental epistemic action is carried out. Peggy attempts to bring to mind a mental image of Alpha Manor; specifically, she tries to retrieve the set of most likely hypotheses about Alpha Manor, as encoded in her generative model during prior experiences of the house. Given the context of Peggy's imagining, it's important that she's aiming to bring to mind the hypotheses her generative model says are *most* likely rather than altering this information in any way. By doing this, Peggy can, instead of physically going to look at Alpha Manor, bring to mind her brain's expectations about what she *would* see were she to actually go look at the house. These are the same hypotheses her brain would initially draw on to construct her perceptual experience were she observing Alpha Manor directly — though in a perceptual case, these hypotheses would, additionally, be confirmed and revised against sensory evidence, which they aren't in a case of imagining. (Of course, Peggy ends up accidentally drawing on the wrong set of hypotheses, hypotheses about Beta Estate; what I have described here is the cognitive process she *tries*, but fails, to perform.)

Once Peggy has formed her mental image and counted the number of windows, she will have eliminated her uncertainty in how to continue conversing with the caller, allowing her to assert her explicit belief

about Alpha Manor's windows. In doing so, she eliminates the prediction error that was initially generated when her generative model said she was responding to the caller, but she was not in fact doing so.

We can see now that both the active and perceptual nature of actuality-oriented imagining are important for understanding Peggy's imagining. Fully understanding her imagining's function requires taking into account the broader context in which it's performed. Peggy is trying to carry out a particular action — making an assertion on the phone — and faces uncertainty about how to do so. This generates prediction error with respect to her brain's hypotheses about her current actions. Her imagining is a mental epistemic action that functions to reduce her uncertainty and therefore to eliminate prediction error. Fully understanding how she carries out this mental epistemic action requires understanding the mechanisms involved in constructing an actuality-oriented image of Alpha Manor — namely, that this is a matter of retrieving a set of information her generative model has collected from prior experiences, in the form of its most likely hypotheses about Alpha Manor.

Now that we have this story about WINDOWS, we can easily tell the same kind of story about DRAPES. In that case, Ed is also in a context of acting: he's at the store carrying out his goal of buying a set of drapes for each of Alpha Manor's windows. However, he faces uncertainty about how to achieve this goal, because he doesn't have an explicit belief about how many windows Alpha Manor has. So, in order to avoid a discrepancy between what his generative model says about how he's currently acting and what his body is actually doing, he must either abandon his goal of buying the drapes or perform some kind of epistemic action to alleviate his uncertainty. Rather than performing a physical epistemic action, such as returning to Alpha Manor and counting its windows, he performs a mental epistemic action, imagining Alpha Manor. The mental epistemic action he tries to perform is basically the same one as Peggy tries: bringing to mind the most likely set of hypotheses about Alpha Manor, given his prior experiences of

it (though, of course, he ends up mis-imagining but in a different way than Peggy does).

The account I just developed describes actuality-oriented imagining as retrieving an existing set of information about the world. This naturally raises the question of whether actuality-oriented imagining is a kind of *remembering* or exactly how it's related to memory. An overarching goal of this paper is to carve off actuality-oriented imagining as a distinctive cognitive capacity; however, if actuality-oriented imagination is really just a kind of memory, this might seem to confound this goal, by implying that our focus should be on studying memory rather than actuality-oriented imagining in isolation. I'll address this concern in the remainder of this subsection.

Perhaps the most obvious place to look when considering this question is *episodic* memory, which involves using mental imagery to recall events that we previously experienced — as when, for example, I remember my tenth birthday party or yesterday's lunch.<sup>12</sup> Actuality-oriented imagining has a lot in common with episodic memory, given its sensory representational format and the fact that it involves bringing to mind imagery of previously experienced objects. Nevertheless, there are some key differences between them. To see these, it's helpful to invoke what Sarah Robins (2020) calls the *target* content of episodic memories.<sup>13</sup> In line with typical definitions of episodic memory in philosophy and psychology, Robins argues that for a mental state to be an instance of episodic remembering, it's necessary that one targets or aims to represent a specific type of content: *particular* events from one's *personal past*, that is, events one personally experienced in the past.

With this in mind, we can see that the target content of actuality-oriented imagining contrasts with episodic memory in two important

12. The term "episodic memory" was introduced in psychology by Tulving (1972) as distinct from *semantic* memory, which I discuss below. The study of episodic memory has recently been taken up by many philosophers of mind (e.g., De Brigard 2014; Michaelian 2016b).
13. Robins notes that she develops the idea of target content from Cummins (1996) but that her application to episodic memory is original.

ways. First, while episodic memory targets the past, actuality-oriented imagining targets the present. In both WINDOWS and DRAPES, for example, the subjects are aiming to imagine Alpha Manor as it is now: they don't currently care how many windows it *had* during some past event but how many windows it *has* now. Second, while episodic memory by definition involves recreating the contents of a prior experience, this isn't true for actuality-oriented imagining: the subjects in our cases might be visualizing Alpha Manor in ways that don't reflect particular past experiences. We can starkly illustrate this difference by considering circumstances under which it would be more efficient to use actuality-oriented imagining than episodic memory to answer the question of how many windows Alpha Manor has. Suppose that Peggy has retained four episodic memories of particular past visits to the house: one in which she's driving past the house and seeing it from the front; one in which she's in the back yard, taking pictures of the back of the house; one in which she's walking from one side of the house around to the back while speaking with a potential buyer; and one in which she's walking from the back, around the other side, to the front. It would be possible for Peggy to answer the question of how many windows are on Alpha Manor by "flipping through" these various episodic memories of specific past visits, since together they include all sides of the house. However, given that the contents of these memories overlap with one another and that they involve irrelevant details (e.g., pulling out her camera to snap photos; speaking with the potential buyer), it would be much more efficient to imagine a single, sequential "tour" of the outside of the house, even if she's never experienced the house this way in real life.

So, actuality-oriented imagination lacks some properties that are fundamental to episodic memory.<sup>14</sup> However, episodic memory isn't

14. I don't mean to take a stance on whether episodic memory and actuality-oriented imagining are *continuous* in the sense of Perrin (2016) and Michaelian (2016a) — that is, whether, besides their distinct types of contents, the two are in some sense the same kind of cognitive process. Some recent work in this area has argued that episodic remembering is continuous in this sense with various kinds of imagery-construction processes, including imagining

the only kind of memory that involves the retrieval of stored information — *semantic* memory also does this. Semantic memory is our capacity to store and retrieve facts and conceptual knowledge in ways not tied to representations of particular events from the personal past. This category is quite diverse. It paradigmatically includes beliefs we have stored in an explicit, linguistic format, such as when retrieving the stored knowledge that Ottawa is the capital of Canada or that Julius Caesar crossed the Rubicon. It's also often taken to include conceptual knowledge that we retrieve using mental imagery (cf. Binder and Desai 2011; Yee et al. 2018), such as when one visualizes a prototypical donkey to bring to mind the proposition that donkeys have manes or visualizes the Italian flag to bring to mind information about which colors are on it. This sort of information is also typically thought to be operative when we deploy stored conceptual knowledge in object recognition. (The role for semantic memory in perception may be even greater in the PP framework I've adopted in this paper, given the way PP says top-down stored information heavily permeates perceptual processing.)

If all of these sorts of processes are included in the category of semantic remembering, it seems like actuality-oriented imagining should be too: if using mental imagery to retrieve general facts is a form of semantic memory, and if perceptually processing particular parts of the world involves drawing on semantic memory, then using mental imagery to retrieve facts about particular parts of the world also sounds like a form of semantic memory. In our WINDOWS and DRAPES cases, this would mean using mental imagery to bring to mind the fact that Alpha Manor has ten windows.

---

the future, imagining counterfactuals, and imagining atemporal scenes (Mullally and Mullally 2013; De Brigard and Gessell 2016; Sant'Anna and Michaelian 2019). If this is right, it could plausibly be extended to actuality-oriented imagining too. Even so, continuists don't tend to deny that, in virtue of their distinct types of content, we can meaningfully refer to memory and imagination as separate capacities (Michaelian 2016b, e.g., refers to episodic memory as a natural kind, despite also thinking it's continuous with future-imagining).

Notice, though, that this makes semantic memory quite a heterogeneous category, especially when we also throw in all the various kinds of non-sensory, linguistic representations we possess (explicit stored knowledge of general facts, of facts about particular parts of the world, of abstract mathematical facts, etc.). Semantic memory thus ends up including any kind of information storage and retrieval that isn't episodic, unlike episodic memory, which is quite narrowly defined in terms of a specific type of target content.<sup>15</sup> This heterogeneity warrants individualized study of the various capacities grouped under the category of semantic memory, in addition to studying the category as a whole. This is in line with existing practices among philosophers and psychologists: although we sometimes study information storage and retrieval at a very general level, we also, for example, study our capacities for object recognition and mathematical learning in isolation from each other. Plausibly, the same should go for actuality-oriented imagination, especially since, as I have already argued and will continue to argue below, it's distinct in various ways from other, similar cognitive capacities.

So, although it may be right to include actuality-oriented imagination alongside various other capacities classified as semantic memory, this doesn't mean that it's not a distinctive cognitive capacity that warrants individualized study.

#### 4. Back to content determination and success conditions

I'll now leverage the previous section's account of actuality-oriented imagining to fill out answers to the content determination and success conditions questions. Specifically, I'll aim to both vindicate and make more precise the intuitions I discussed in §2. There I argued that something like the following two conditions seem to distinguish actuality-oriented imagining from perception and other kinds of imagining:

15. Besides semantic and episodic memory, there's also *nondeclarative* memory, which includes skills, habits, and classical conditioning. However, as Michaelian (2016b, sec. 2.6) argues, nondeclarative memory plausibly doesn't store and retrieve representations. As such, it's outside the scope of my discussion here.

(1) On content determination: imagining some object requires drawing on stored information about that object to form one's mental image.

(2) On success conditions: to be successful, an actuality-oriented imagining need only veridically represent its object in respects relevant to the question(s) a subject aims to answer.

I'll take each of these points in turn.

To see why something like (1) falls out of §3's PP-based account, we have to consider the overall cognitive process that this account says is involved in forming an actuality-oriented mental image. Actuality-oriented imagining is a mental epistemic action meant to reduce uncertainty about achieving one's current goals, in cases where this requires one to form a belief about the actual world. Given this, it involves a subject trying to bring to mind true information about the world. Consideration of the perceptual mechanisms involved in actuality-oriented imagining tells us more specifically how a subject goes about trying to do this — namely, by trying to bring to mind the set of perceptual hypotheses about the object she's imagining that her brain's generative model says are most likely. So, rather than trying to form a relevant belief by somehow discovering *new* information about the world, a subject is using the imagination as a means of trying to bring to mind an *existing* set of information.

This account verifies the intuition that Peggy fails to imagine Alpha Manor in WINDOWS. Peggy is trying to bring to mind the set of most likely hypotheses about Alpha Manor, but what she actually ends up doing is bringing to mind a set of hypotheses about Beta Estate. It's not that she brings to mind some false information about Alpha Manor, or that she doesn't possess any stored information about what Alpha Manor looks like so has to make a best guess. Rather, she retrieves the wrong information altogether, information about a totally different object. This means that Peggy fails to carry out the cognitive process constitutive of actuality-oriented imagining Alpha Manor. In

other words, if trying to imagine Alpha Manor in an actuality-oriented case is equivalent to trying to bring to mind hypotheses about Alpha Manor, then Peggy altogether fails to imagine Alpha Manor when she brings to mind hypotheses about Beta Estate.

This account thus both vindicates and makes more precise intuition (1). The intuition is correct because imagining some object in an actuality-oriented case involves trying to retrieve stored information about that object; one must actually retrieve such information if one is to carry out the process one is trying to implement, and one fails altogether to do so when one brings to mind information about a totally different object. And we can make this point more precise by explaining what it means to imaginatively bring to mind stored information in such cases: the cognitive process involved is one of trying to bring to mind the set of perceptual hypotheses about an imagined object that the brain's generative model says are most likely.

There's a sense in which what I just argued seems to merely pass the buck. We were asking how an actuality-oriented mental image gets to be about some object, and my response referred us to facts concerning what the stored information used to construct an image is about. The obvious question this raises is how that stored information itself comes to be about the object. However, this is a large, general question about the semantics of stored mental content, one that's distinct from my focus on the nature of actuality-oriented imagining. The question I addressed in this section is: What kind of process is one trying to implement when trying to actuality-oriented imagine some object, such that carrying out this process is necessary for bringing to mind a mental image of that object? My response is that because actuality-oriented imagining is a process aimed at bringing to mind stored information about a particular object, one fails to imagine that object when one draws on stored information about a different object. A *distinct* question asks how that stored information's content was originally determined, prior to the time that it goes on to determine the content of a conscious mental image; plausibly, these questions can be kept apart. Robins (2016) makes a similar point in her discussion of

episodic memory. She argues that the question of how memory traces (i.e., episodic representations that are stored in memory and apt to be activated to cause episodes of episodic remembering) are stored in relation to one another and activated to produce conscious mental imagery is distinct from the question of how those stored traces themselves get their contents. The former question is analogous to my concerns in this paper and is relevant to determining what a conscious mental image, which is constructed by drawing on stored information, represents; the latter is a much larger issue about how stored mental representations acquire their content.

Given my adoption of the PP framework in this paper, I can at least make this question more precise and gesture at a possible answer. Given how PP conceptualizes all stored information about the world as a generative model consisting in probability distributions over hypotheses about the states of the world, the relevant, more precise question here is: How does such a generative model come to represent the world? Recent work in the PP framework has argued that the generative model represents by *resembling* the causal-probabilistic structure and dynamics of the world (Gładziejewski 2016; Wiese 2017; Williams 2018). I think this answer may need to be supplemented by appealing to a *causal* condition as well, given that the generative model acquires its resemblance to the world via perceptual causal connections that allow it to learn the world's structure and dynamics (i.e., it's not *mere* resemblance that grounds representation but resemblance acquired via appropriate causal connection to the world). In any case, since this is a very large topic that has been considered by others in much detail, I'll now set it aside.

Returning to our consideration of (1): I said in §2.1 that (1) doesn't seem to apply to perception nor to other kinds of imagining. The present framework helps to explain this too.

Although in PP perceptual experience also involves drawing on the brain's best hypotheses about the external world, perception is a process in which these hypotheses are also confirmed by, or revised against, sensory evidence. In a case where one perceives some part

of the world one is already familiar with, it's likely that information one already possesses about that part of the world will feature in the construction of one's experience, as it does in actuality-oriented imagining. However, it's not *necessary*, to perceive some object, that one possesses previously stored information about it. One can perceive novel objects by updating one's current perceptual hypotheses based on sensory evidence that one is perceiving something novel. In other words, it can be the case that, in the very process of acquiring information about an object for the first time, one comes to perceive that object.

As I said in §3.1, it's not immediately obvious how to explain, within the PP framework, the computations involved in non-actuality-oriented imaginings to which (1) does not seem intuitively to apply. However, whatever account we end up giving, it seems clear that those sorts of imagining will involve a process that's different from what PP says about actuality-oriented imagining. As per my discussion in §2.1, it's plausible that my mental image can represent Socrates when I imagine an arbitrary man and stipulate that I'm imagining Socrates. In that case, it's clearly not true that I need to bring to mind hypotheses made most likely by my previous experiences of Socrates in order to imagine him, because I don't have any such stored information. This suggests that, whatever exactly the imagery-construction process involves in this kind of case, it will be a different kind of cognitive process than the one involved in actuality-oriented imagining. So, if there's any sense in which it's possible to try, and fail, to imagine Socrates in this kind of case (although such failure may not even be possible — cf. Dorsch 2012, ch. 3), it will be a different kind of failure from Peggy's failure when she tries, and fails, to imagine Alpha Manor. If I don't need to possess any prior information about Socrates to imagine him in this stipulative way, then I can't fail to imagine him by failing to draw on the right information.

Turn now to (2). We can break this intuition down into two sub-intuitions. First, to be successful, an actuality-oriented imagining must veridically represent the properties of an object that are relevant to



answering the question(s) a subject is aiming to answer. Second, to be successful, an actuality-oriented imagining need not veridically represent the properties of an object that aren't relevant to a subject's question(s). We can vindicate and precisify both of these sub-intuitions by again appealing to §3's account of actuality-oriented imagining.

As per §2.2, we can derive a mental capacity's success conditions from the conditions under which it fulfills its cognitive function. According to §3's account, the function of an actuality-oriented imagining is to reduce uncertainty about how to continue acting, thus eliminating prediction error arising from the fact that, while one's generative model says that one is acting, one's actions have halted in the face of uncertainty. For Ed in DRAPES, for example, this means reducing his uncertainty about how to continue buying new drapes for Alpha Manor's windows.

Now, at the time of this case, it will *seem* to Ed that his imagining has fulfilled this function, since it removes his feelings of uncertainty. When he imagines Alpha Manor with the wrong number of windows, he won't immediately be aware of any discrepancies between how he takes the world to be and how the world really is, because he will continue to act in a way that seems to conform to what his generative model says he's doing — that is, buying drapes for each of Alpha Manor's windows. However, it's important that the predictive brain aims not just at minimizing *apparent* discrepancies, in the short term, between its model and the world; rather, its aim is to minimize *actual* discrepancies. That's because merely doing the former isn't conducive to minimizing overall prediction error in the long run — the more actual differences there are between how I think the world is and how the world actually is, the more likely it is that I'll falsely predict the states of my environment at some point and therefore encounter an unexpected or surprising situation. We can see this clearly in DRAPES. At the time of the case, Ed's generative model says that he's buying new drapes for each of Alpha Manor's windows, and it seems that he's successfully doing so. What he's really doing, though, is buying one too many sets of drapes. This discrepancy will result in prediction error in

the near future. When Ed is about to arrive home, he will predict that he has exactly one set of drapes for each window, that he'll go on to install those drapes accordingly, and so on. Once he arrives home, it will become apparent that many of his expectations are false, generating prediction error.

So, the function of Ed's imagining is to reduce his uncertainty in a way that contributes to *successfully* continuing to carry out his goal of buying drapes for each window. This function sets his imagining's success conditions. Ed's mental image of Alpha Manor must have the correct number of windows to be successful; too many or too few will result in a discrepancy between the predictions of Ed's generative model and how the world actually is. At the same time, it also wouldn't prevent Ed's imagining from fulfilling its function if he inaccurately imagines other properties of Alpha Manor. Those other properties aren't currently relevant to minimizing prediction error, since he isn't basing new beliefs on them or acting on the basis of them. Once he forms his belief about the number of windows, his mental image, with whatever other inaccuracies it contains, serves no further purpose.

So, the PP-based account of actuality-oriented imagining helps us vindicate (2): this account implies that only the properties of a mental image relevant to the brain's core function of prediction error minimization must be veridical, and these properties line up with our original intuitions about which properties must be veridical. We can also appeal further to what this account says about actuality-oriented imagining's perceptual aspects to make this intuition more precise. An actuality-oriented mental image is constructed from the set of perceptual hypotheses one's generative model says are most likely. We can unpack (2) in terms of which perceptual hypotheses used to generate a mental image must be true in order for the imagining to be successful — that is, for the imagining to be conducive to successful prediction error minimization. To be successful, an actuality-oriented imagining need not be constructed from a set of hypotheses that are all true; rather, only the subset of hypotheses that are relevant to prediction

error minimization need be true, while those that aren't relevant to this function need not be.<sup>16</sup>

The PP framework also helps us see why this is a fact that's distinctive of actuality-oriented imagining, in comparison with perception and other kinds of imagining.

It might seem at first that the arguments I just gave about actuality-oriented imagining extend to the success conditions of perceptual experience, in that only the elements of a perceptual experience upon which we're currently basing explicit beliefs need to be accurate for an experience to count as successful. However, the PP framework for perception actually predicts the opposite of this. On the PP account, all aspects of perceptual representations are produced in a process of the brain trying to minimize prediction error, by confirming its perceptual hypotheses against sensory inputs. This process aims to eliminate discrepancies between my brain's current predictions and the way the world is, using sensory evidence to try to correct such discrepancies. So, any perceptual experience that mis-represents the world is one that has failed to reduce discrepancies between the brain's hypotheses and what's true.<sup>17</sup> Contrast this with actuality-oriented imagining cases, in which a subject is retrieving information about the world in a process wherein this information is not automatically tested against sensory input. Here if some details of the resulting mental image are

16. At the end of §3, I distinguished actuality-oriented imagining from episodic memory based on their respective target contents. On some views of memory, my arguments in this section about success conditions also constitute a difference. Robins (2020), for example, takes any deviation from a remembered experience's contents to be memory errors, a stricter condition than what I just defended for actuality-oriented imagining (and one that seems to fit with our intuitive conception of episodic memory's success conditions). However, this view isn't shared by all — De Brigard (2014) and Michaelian (2016b), for example, deny it.

17. This is an oversimplification in that — as I mentioned above in note 9 — it glosses over the role of precision weighting and attention. In brief, the brain may settle into less precise interpretations of parts of the world to which we're not perceptually attending; in those cases, our experience may not precisely or fully determinately represent what's in our perceptible environment. In such cases, the brain also won't revise its model of some imprecisely-represented part of the world to be more precise than its perceptual representation.

inaccurate, it won't be as a result of that subject failing to properly revise some hypothesis against sensory input, since that's not the kind of process she's implementing when she imagines. The inaccuracies in her mental image may be grounded in some inaccuracies in the stored information she's drawing on, which could have gotten there either during or after the process of originally acquiring this information during past perceptual experience. But that just means there's a deficiency in her stored information about the world, not a deficiency with respect to whether the mental image fulfills its function.

The same success conditions also wouldn't apply to other cases of imagining. Often we don't base new beliefs about the world on our imaginings, such as when merely fantasizing; in such cases, one can't be susceptible to the kind of failure to which Ed is susceptible.<sup>18</sup> What about imaginings on which we base beliefs about what's possible? Beliefs about possibility and/or impossibility can be true or false and also may help to shape our future predictions about the world. They may do this by, for example, helping to narrow the space of hypotheses the brain takes into account when deciding which is most likely. So, such beliefs, if false, plausibly could negatively impact long-run prediction error minimization, and thus so can the imaginings we use to form such beliefs. Nevertheless, if the function of such imaginings is to tell us about which hypotheses are possible, achieving this function doesn't require veridically representing the object one imagines in any particular respects. Rather, it just requires accurately representing what's possible. So, relative to this function, we would expect success conditions that are distinct from those for actuality-oriented imagining, whose function requires veridically representing an object in some relevant respect(s).

18. It's a tricky question how to fit phenomena like fantasy and fictional engagement into the PP framework, since it's not clear exactly how they're relevant to minimizing prediction error. However, this problem generalizes to other accounts of cognition besides PP — it's just generally difficult to figure out why we have the capacity to fantasize and think about fictional worlds, given that it's not obvious exactly what evolutionary advantage this confers.

## 5. Conclusion

I have tried, in this paper, to carve off actuality-oriented imaging as distinctive in both its metaphysics of representation and the success conditions governing it. Its distinctiveness is grounded in the fact that it has characteristics that are both perceptual and active, aspects that this paper attempted to unify into a single account. In conclusion, I'll sketch two implications of what I have argued in this paper for the philosophy of imagination more generally.

First, my arguments that actuality-oriented imagining is importantly distinct from perception and other kinds of imagining have implications for how we conceptualize the different "sub-categories" of imagination. As is often noted, there's a huge range of domains in which we employ what seem aptly described as "imaginative" capacities. Many of these have generated their own focused research programs: there are large literatures on the use of imagination in modal epistemology (e.g., Yablo 1993), mindreading (e.g., Goldman 2006), and engagement with fiction (e.g., Walton 1990), to name only a few. The arguments of this paper suggest that actuality-oriented imagining deserves a place alongside these other uses of imagination, as its own sub-category. This is especially true given that understanding all the diverse ways we use the imagination seems like an important part of evaluating the prospects of systematizing these ways into some kind of unified framework (for further discussion of such prospects, see Dorsch 2012; Kind 2013; Abraham 2016).

Second, getting clear on the nature of actuality-oriented imagining, as I unpacked it in this paper, has implications for how we conceptualize the question at the heart of the epistemology of imagination. Some recent philosophy takes it that the imagination is by default a faculty unconstrained from reality, an important philosophical puzzle being how we're sometimes able to "rein in" this faculty for epistemic purposes, such as thinking about what's possible (cf. Kind and Kung 2016). However, many cases of actuality-oriented imagining seem, on reflection, to be rather mundane yet epistemically significant — it's

quite natural and effortless to use mental imagery as a means of forming a belief about the number of windows on one's house, for example. Furthermore, such uses of the imagination are, as I remarked in §3.1, rather computationally simple: they involve bringing to mind existing stored information about the world without transforming or recombining the information in any way. And they're implemented using perceptual cognitive mechanisms that — assuming the PP framework is on the right track — are generally in the business of drawing on existing information about the world to construct representations of actuality. In light of all this, the real puzzle seems not to be how we can use such a reality-transcendent faculty for epistemically useful purposes. Instead, the more puzzling question seems to be how cognitive mechanisms so suited to representing actuality are also able to so easily transcend it. Rather than taking reality-transcendent uses of imagination as basic and asking how it's possible to bring the imagination in line with reality, it may be more fruitful to start from actuality-oriented uses and ask how it's possible to go beyond these.<sup>19</sup>

## References

- Abraham, Anna. 2016. "The Imaginative Mind." *Human Brain Mapping* 37 (11): 4197–4211.
- Balcerak Jackson, Magdalena. 2016. "On the Epistemic Value of Imagining, Supposing, and Conceiving." In *Knowledge Through Imagination*, edited by Amy Kind and Peter Kung, 41–60. Oxford: Oxford University Press.
- Balcerak Jackson, Magdalena. 2018. "Justification by Imagination." In *Perceptual Imagination and Perceptual Memory*, edited by Fiona Macpherson and Fabian Dorsch, 209–226. Oxford: Oxford University Press.
19. For feedback on earlier drafts of this paper, I'm grateful to David Barnett, Jennifer Nagel, Madeleine Ransom, Gurpreet Rattan, Margot Strohminger, an anonymous referee for this journal, and audiences at the Canadian Philosophical Association and the University of Düsseldorf. This research was supported by the Social Sciences and Humanities Research Council of Canada.

- Binder, Jeffrey R., and Rutvik H. Desai. 2011. "The Neurobiology of Semantic Memory." *Trends in Cognitive Science* 15 (11): 527–536.
- Clark, Andy. 2016. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press.
- Cummins, Robert. 1996. *Representations, Targets, and Attitudes*. Cambridge, MA: MIT Press.
- De Brigard, Felipe. 2014. "Is Memory for Remembering? Recollection as a Form of Episodic Hypothetical Thinking." *Synthese* 191 (2): 155–185.
- De Brigard, Felipe, and Bryce S. Gessell. 2016. "Time Is Not of the Essence: Understanding the Neural Correlates of Mental Time Travel." In *Seeing the Future: Theoretical Perspectives on Future-Oriented Mental Time Travel*, edited by Kourken Michaelian, Stanley B. Klein, and Karl K. Szpunar, 153–179. Oxford: Oxford University Press.
- Dewhurst, Joe. 2017. "Folk Psychology and the Bayesian Brain." In *Philosophy and Predictive Processing*, edited by Thomas Metzinger and Wanja Wiese. Frankfurt: MIND Group.
- Dorsch, Fabian. 2012. *The Unity of Imagining*. Berlin: De Gruyter.
- Drayson, Zoe. 2017. "Modularity and the Predictive Mind." In *Philosophy and Predictive Processing*, edited by Thomas Metzinger and Wanja Wiese. Frankfurt: MIND Group.
- Friston, Karl, Thomas FitzGerald, Francesco Rigoli, Philipp Schwartenbeck, John O'Doherty, and Giovanni Pezzulo. 2016. "Active Inference and Learning." *Neuroscience and Biobehavioral Reviews* 68: 862–879.
- Friston, Karl, Francesco Rigoli, Dimitri Ognibene, Christoph Mathys, Thomas FitzGerald, and Giovanni Pezzulo. 2015. "Active Inference and Epistemic Value." *Cognitive Neuroscience* 6 (4): 187–214.
- Gładziejewski, Paweł. 2016. "Predictive Coding and Representationalism." *Synthese* 193 (2): 559–582.
- Goldman, Alvin I. 2006. *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford: Oxford University Press.
- Graham, Peter J. 2014. "The Function of Perception." In *Virtue Scientia: Bridges between Virtue Epistemology and Philosophy of Science*, edited by Abrol Fairweather, 13–31. Dordrecht: Synthese Library.
- Hohwy, Jakob. 2013. *The Predictive Mind*. Oxford: Oxford University Press.
- Kind, Amy. 2001. "Putting the Image Back in Imagination." *Philosophy and Phenomenological Research* 62 (1): 85–109.
- Kind, Amy. 2013. "The Heterogeneity of the Imagination." *Erkenntnis* 78 (1): 141–159.
- Kind, Amy. 2018. "How Imagination Gives Rise to Knowledge." In *Perceptual Imagination and Perceptual Memory*, edited by Fiona Macpherson and Fabian Dorsch, 227–246. Oxford: Oxford University Press.
- Kind, Amy. 2019. "Mary's Powers of Imagination." In *The Knowledge Argument*, edited by Sam Coleman, 161–179. Cambridge: Cambridge University Press.
- Kind, Amy, and Peter Kung. 2016. "Introduction: The Puzzle of Imaginative Use." In *Knowledge Through Imagination*, edited by Amy Kind and Peter Kung, 1–37. Oxford: Oxford University Press.
- Kirchhoff, Michael D. 2018. "Predictive Processing, Perceiving and Imagining: Is to Perceive to Imagine, or Something Close to It?" *Philosophical Studies* 175: 751–767.
- Kung, Peter. 2016. "You Really Do Imagine It: Against Error Theories of Imagination." *Noûs* 50 (1): 90–120.
- Maguire, Eleanor A., and Sinéad L. Mullally. 2013. "The Hippocampus: A Manifesto for Change." *Journal of Experimental Psychology: General* 142 (4): 1180–1189.
- McGinn, Colin. 2004. *Mindsight: Image, Dream, Meaning*. Cambridge, MA: Harvard University Press.
- Macpherson, Fiona. 2017. "The Relationship between Cognitive Penetration and Predictive Coding." *Consciousness and Cognition* 47: 6–16.
- Metzinger, Thomas. 2017. "The Problem of Mental Action: Predictive Control without Sensory Sheets." In *Philosophy and Predictive Processing*, edited by Thomas Metzinger and Wanja Wiese. Frankfurt: MIND Group.

- Michaelian, Kourken. 2016a. "Against Discontinuism: Mental Time Travel and Our Knowledge of Past and Future Events." In *Seeing the Future: Theoretical Perspectives on Future-Oriented Mental Time Travel*, edited by Kourken Michaelian, Stanley B. Klein, and Karl K. Szpunar, 62–92. Oxford: Oxford University Press.
- Michaelian, Kourken. 2016b. *Mental Time Travel: Episodic Memory and Our Knowledge of the Personal Past*. Cambridge, MA: MIT Press.
- Perrin, Denis. 2016. "Asymmetries in Subjective Time." In *Seeing the Future: Theoretical Perspectives on Future-Oriented Mental Time Travel*, edited by Kourken Michaelian, Stanley B. Klein, and Karl K. Szpunar, 39–61. Oxford: Oxford University Press.
- Pezzulo, Giovanni. 2017. "Tracing the Roots of Cognition in Predictive Processing." In *Philosophy and Predictive Processing*, edited by Thomas Metzinger and Wanja Wiese. Frankfurt: MIND Group.
- Pezzulo, Giovanni, Emilio Cartoni, Francesco Rigoli, Léo Pio-Lopez, and Karl Friston. 2016. "Active Inference, Epistemic Value, and Vicarious Trial and Error." *Learning and Memory* 23 (7): 322–338.
- Pezzulo, Giovanni, and Stefano Nolfi. 2019. "Making the Environment an Informative Place: A Conceptual Analysis of Epistemic Policies and Sensorimotor Control." *Entropy* 21 (4).
- Robins, Sarah. 2016. "Representing the Past: Memory Traces and the Causal Theory of Memory." *Philosophical Studies* 173 (11): 2993–3013.
- Robins, Sarah. 2020. "Mnemonic Confabulation." *Topoi* 39: 121–132.
- Sant'Anna, André, and Kourken Michaelian. 2019. "Thinking about Events: A Pragmatist Account of the Objects of Episodic Hypothetical Thought." *Review of Philosophy and Psychology* 10 (1): 187–217.
- Stivers, Tanya. 2010. "An Overview of the Question-Response System in American English Conversation." *Journal of Pragmatics* 42 (10): 2772–2781.
- Tidman, Paul. 1994. "Conceivability as a Test for Possibility." *American Philosophical Quarterly* 31 (4): 297–309.
- Tulving, Endel. 1972. "Episodic and Semantic Memory." In *Organization of Memory*, edited by Endel Tulving and Wayne Donaldson, 381–403. New York: Academic.
- Walton, Kendall L. 1990. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, MA: Harvard University Press.
- Wiese, Wanja. 2017. "What Are the Contents of Representations in Predictive Processing?" *Phenomenology and the Cognitive Sciences* 16 (4): 715–736.
- Williams, Daniel. 2018. "Predictive Processing and the Representation Wars." *Minds and Machines* 28 (1): 141–172.
- Williamson, Timothy. 2016. "Knowing by Imagining." In *Knowledge Through Imagination*, edited by Amy Kind and Peter Kung, 113–123. Oxford: Oxford University Press.
- Yablo, Stephen. 1993. "Is Conceivability a Guide to Possibility?" *Philosophy and Phenomenological Research* 53 (1): 1–42.
- Yee, Eiling, Michael N. Jones, and Ken McRae. 2018. "Semantic Memory." In *Language and Thought*, edited by Sharon L. Thompson-Schill, 319–356. Vol. 3 of *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, 4th ed., edited by John T. Wixted. New York: Wiley.