7 From Prediction to Imagination

Max Jones and Sam Wilkinson

There is an increasingly popular framework for thinking about what the brain is fundamentally in the business of doing. We call this the predictive processing framework (PPF), and it takes the brain to be a hierarchically arranged prediction machine (Clark, 2013, 2016; Hohwy, 2013). In this chapter, we explore the consequences that this framework has for our understanding of imagination. In other words, we ask: "If you buy into this framework, what account of imagination should you give?"

In one sense, the PPF can be seen to demystify the imagination. One of the seemingly anomalous features of the imagination, from a naïve perspective, is its capacity for endogenously and creatively generating content, for conjuring up worlds in the mind. Yet, in the PPF approach, our basic worldly interactions through perception and action are already essentially dependent on imagination-like generative processes. However, the PPF's commitment to the ubiquity of imagination-like processes in our more mundane worldly interactions may also give rise to problems in identifying and explaining imagination as a distinct process in its own right. In particular, the PPF may struggle to explain the distinctive imaginative capacity for deliberately departing from immediate reality in a constrained and purposeful manner.

We proceed as follows. We start by introducing the PPF (section 1: Introducing the Predictive Processing Framework). Arguably, the PPF is most intuitively appealing as an account of perception and perceptual cognition, but it has also been extended to an understanding of action, through the notion of active inference. We then (in section 2: Predictive Processing and Imagination) present a first attempt at accounting for imagination within the PPF in terms of perceptual inference, which has been assumed by several theorists to be rather straightforward. We then (in section 3: Imaginative Agency and Imaginative Constraints) argue that this first attempt faces problems in accounting for the purposeful and constrained nature of deliberate imagination, which sets it aside from processes like dreaming, hallucination, or mind-wandering. We then consider (in section 4: Imagination as Mental Action) an alternative PPF account of imagination that focuses more on active inference and takes imagination to be based on action simulation. Although this account is more promising, it fails, on its own, to fully account for more drastic imaginative departures from reality. With this in mind, we gesture toward a possible way forward that focuses on the role of language in cueing and shaping imagination.

94

Introducing the Predictive Processing Framework (PPF)

The PPF is introduced and explained in a number of different ways by different people (see Clark, 2013 for a classic formulation). Our favorite way is to think about two things. The first is the *ambiguity* inherent in the information that we get from the world (including, crucially, ourselves as part of that world). The second is the *efficiency* with which the brain ought to operate.[1] Let's examine these two things in turn.

Bayesian Resolution of Ambiguity

The world is a noisy and *ambiguous* place. For any system trying to make sense of the world, which we can gloss roughly as trying to come up with hypotheses about the world, there is always more than one hypothesis compatible with any given piece of data. How, then, in the face of this ambiguity, does any such system pick a hypothesis out of those with equally good fit? According to the PPF the answer to this question is the same as the one that Thomas Bayes gave to it more than 250 years ago (Bayes, 1763): The system needs to use background knowledge, namely, knowledge about the probabilities of each hypothesis independently of (prior to) the evidence (the data). This is called the "prior probability," or "prior" for short, and it serves to bias hypothesis selection in the face of equal (or near-equal) fit.[2] A consequence of this idea is that a hypothesis can be selected even though it has relatively bad fit, as long as it has a high enough prior probability. Conversely, a hypothesis with a relatively good fit might not be selected if its prior probability is too low.

Let us illustrate this Bayesian story with an example (borrowed from Pezzulo, 2014). Suppose you are in a bedroom, and you hear a downstairs window creak. And suppose for simplicity's sake that two hypotheses present themselves: "That's the wind blowing the window" or "That's a burglar climbing in through the window." Both have adequate fit in the sense that, if they were true, that is the sound that might plausibly be produced. However, the probability of each hypothesis independently of the squeak (viz. the prior probability) might be quite different. Suppose you know that you live in a neighborhood with a very low crime rate, and you know that it's a windy night. That would give the wind hypothesis a higher prior probability, and that is the hypothesis that ought to be selected.

At this point it is crucial to emphasize that this example involves a person (you) dealing with experiential information (the window squeak) in a Bayesian manner. The PPF intends this to apply to even the lowest levels of neural inference. To use this same example, by the time that you have experienced the qualitative features of

---

1 By "efficiency" we mean an optimal trade-off between speed, accuracy, and energy expenditure. By "ought to" we don't mean anything robustly normative, but rather an expectation that evolutionary pressures will lead to efficient systems.

2 Where these priors come from is theoretically up for grabs. What we know is that at least some are the result of statistical learning from past experience (exposure to the world). Some might be innate in some sense (whether what counts as innate is the priors themselves or the propensity to acquire certain priors).

a squeak (even prior to recognizing it as a *window* squeak) your nervous system has already made countless earlier Bayesian inferences based on the raw (preconscious, pre-experiential) sensory input.

Efficient Neural Implementation Through Predictive Processing

According to the PPF, this is the Bayesian strategy that the brain adopts in the face of an ambiguous world. But how is it implemented in the brain? The answer here is: in a way that is *efficient*, which is to say, both *anticipatory* and *energy-saving*. Given that the human nervous system has a wealth of background knowledge that it brings to bear on an ambiguous world, we can think of each encounter with the world as already being anticipatory. As energy impacts on your sensory surfaces, it is already being greeted by expectations on the part of the system. This, one might argue, is simply a physiological extrapolation of the notion of Bayesian priors. It also paves the way for a tremendously energy-saving information-processing strategy. To see why, think about the fundamental principle underlying data compression. When passing a message between a sender-receiver pair, you optimize bandwidth by leaving out the message that the receiver can fill in for itself. This principle translates to the human brain in that the relevant senders and receivers are parts of the cortical hierarchy. In perceptual processing, say, visual processing, as light hits the retina, it is already being greeted by expectations. But what gets passed up the processing chain? The part of the signal that hasn't already been predicted, namely, prediction error (Feldman and Friston, 2010). What we end up with is a rich but hugely efficient and proactive system in which only newsworthy aspects of the nervous system are explicitly neurally represented. Furthermore, what you experience at any given time is not primarily determined by sensory input, but rather by the hypothesis that your brain has adopted to best predict (explain away/quash) that sensory input.

Two Tweaks: Hierarchy and Precision

We have in place the basic picture: the Bayesian strategy, and its neural implementation through prediction error minimization. However, there are two crucial tweaks to this picture. The first is the notion of a hierarchy of predictive hypotheses; the second, the notion of precision.

According to the PPF, predictive hypotheses are *hierarchically organized*, with the hypotheses of one level feeding into the level below (and prediction error feeding into the level above). "Higher" parts of the hierarchy are, roughly, those parts that are further away from the sensory stimulus. These tend to be at slower temporal time-scales, and a higher level of abstraction. At the very top they might correspond to relatively stable beliefs. "Lower" parts of the hierarchy are closer to the sensory stimulus. These tend to be at faster temporal frequencies, and at low levels of abstraction (i.e. concrete and specific). These, for example, might correspond to early stages of visual processing: your brain's early statistically driven attempts to make sense of noisy inputs. Of course, in order to express these neurally encoded predictions we need to use rough-and-ready descriptions in natural language, but

there is nothing linguistic about the priors/hypotheses (e.g. "light tends to come from above" / "This is a face") themselves.

Let us return to the case of the squeaking window. At the stage where those two hypotheses (burglar vs. wind) are competing, a great deal of ambiguity has already been resolved, in a Bayesian fashion, at lower levels of the hierarchy. For example, in early stages of auditory processing, the qualities of the sound will have been settled upon, giving rise to the conscious experience being a certain way, qualitatively speaking. Higher up the hierarchy, that sound gets interpreted as a creaking window, as opposed to, say, a screeching cat. The direction of causation is from the (events represented in the) lower regions of the hierarchy, to the (events represented in the) higher regions of the hierarchy. However, the direction of the Bayesian inference is from the effects to the causes.

So much for *hierarchy*, now for *precision*. Although incoming signals are always ambiguous, in different contexts the degree of ambiguity will differ. To maximize its predictive success, the brain needs to accurately estimate how much ambiguity (uncertainty) there will be. In other words, it needs to make second-order predictions, namely, predictions about how much it should rely on its predictions (which amounts to how much it should pay heed to the prediction-error).

In contexts where low ambiguity is expected (high signal-to-noise ratio), higher precision will be demanded, and the prediction error will be taken more seriously. Conversely, when there is high ambiguity (or, through some top-down influence, there is no interest in the stimulus), low precision is demanded, and the prediction error will be taken less seriously. This is called "precision weighting" (Hohwy, 2013), it amounts to adjusting "the gain" on prediction error, and is taken to be modulated by neurotransmitters such as dopamine (Corlett et al., 2010).

It has been hypothesized that the turning up of precision is the mechanism that underlies attentional focus. Take visual attention as an example. If you are attending to a particular part of your visual field, you have turned the precision right up on your perceptual hypothesis. The result is more precise and informationally rich, but also (other things being equal) higher-risk in terms of error. Thankfully, in appropriate contexts, say, good lighting conditions, you can afford to take that risk because the environment is informationally reliable. If, however, you are walking through the woods at dusk, excessive levels of precision-weighting might lead you to mistake tree-stumps for lurking individuals.

The turning down of precision is also seen as equally important. It corresponds to perceptual decoupling from the world; it allows your brain to hold on to hypotheses that do not adequately predict the world (since the prediction error is down-modulated to the point of being ignored). In the context of *active inference*, which we are about to introduce now, the turning down of precision is central to the inhibition of overt action.

From Perception to Action

So the main business of brains like ours, according to the PPF, is the minimization of errors in the prediction of sensory inputs. However, the minimization of prediction

error isn't only achieved by the brain updating its predictions about the world (which results in perception and belief), it is sometimes achieved by bringing the world, usually the body, in line with the predictions (Clark, 2016: Chapter 4; Feldman and Friston, 2010). The result of this is bodily action. To use a specific example, when I see someone with their hand up, my brain has adopted the hypothesis that someone has their hand up: My brain responds to the world. However, when I move to raise my hand, my nervous system adopts the proprioceptive prediction that my hand is raised. That prediction is then fulfilled by my subsequent hand-raising. Thus an organism's capacity for endogenous bodily movement is just another manifestation of prediction error minimization. It just happens that the relevant prediction in the case of action is a self-fulfilling prophecy.

This distinction between the updating of the hypothesis to fit the world and the updating of the world to fit the hypothesis amounts to the distinction between *perceptual inference* and *active inference*. Although in theory these are two quite distinct manifestations of the predictive processing apparatus, in practice organisms are constantly doing a bit of both. For example, when you are looking through a rainy windscreen, you might perform exploratory movements from side to side to see what is beyond the droplet-covered glass. Here your nervous system is effortlessly toggling between active and perceptual inference to generate an optimal hypothesis of what's in the world, including the intricate subtleties of your place in it. The central mechanism behind this toggling is precision. Precision is down-modulated during the formation of the proprioceptive prediction (the functional equivalent of what in more classical architectures is known as the motor command), otherwise it couldn't be formed at all (since it would be immediately updated to fit the world) and then increased in such a way as to bring the body in line with it.

Predictive Processing and Imagination

How does the PPF accommodate imagination? What is going on in our predictive brains when we imagine something? Here we go through a number of options.

Imagination as the Fundamental Building Block of Experience

Several theorists have taken this to be rather unproblematic, seeing the PPF as especially suited to accounting for imagination. For example, Andy Clark states that the PPF "means that perception (at least, as it occurs in creatures like us), is co-emergent with (something quite like) imagination" (Clark, 2015: 26). And as Michael Kirchhoff (Kirchhoff, 2018: 752) puts it, there is a "deep unity" in the idea, central to the PPF, that "perception and imagination are the psychological results of the brain generating its own sensory inputs top-down." This notion of the "brain generating its own input top-down," though striking, is perhaps a little misleading, at least insofar as it holds on to the idea that what we experience, our phenomenology, at a given time is determined or constituted by "input." A more

complete embracing of the PPF and an understanding of how radically it departs from predominantly bottom-up views leaves us with the notion that our phenomenology at any given time is not determined – still less constituted by – an input (whether self-generated or not) that predates (even fractionally) the experience. Instead, the story goes, it is a forward-looking projection, leaving the role of input as one of correction and constraint: The difference between the imagining mind (or indeed the hallucinating mind) and the perceiving mind is much less great than in standard views, since the role of in-the-moment environmental information is much smaller. What is needed in order to take cognition offline, to break the link with the world, is to turn down the precision in a way that makes model-building (often deliberately, in the case of imagination, and accidentally in the case of hallucination) unconstrained by, and unresponsive to, *sensory input*.

In a sense, this amounts to a revisionist metaphysics of perceptual experience. Contrary to what we might think, perception *actually* involves imagination. On this view, the oft-cited dictum that "perception is controlled hallucination" is almost right. But the word "hallucination" already implies that there is no correspondence to reality. What works better is: "Perception is constrained imagination." Hallucination, on the other hand, is imagination unconstrained by sensory input. This construes imagination as a fundamental building block of PPF: the raw material of all experience, whether offline or online.

Imagination as (a Subspecies of) Offline Cognition

This interpretation of the relationship between the PPF and imagination effectively equates the mind's selected predictive hypothesis (which determines the experience at any given time) with imagination. On this view, any conscious experience – perception, hallucination, dreaming etc. – involves imagination. But one might argue that this seems like an unorthodox terminological stipulation, and an unhelpful one at that. Isn't it more apt to say that the online version of this should be called perception and the offline version (or a subset thereof at least) called imagination? If this is indeed the case, one might think (*pace* Clark) that there is no reason that they would co-emerge. Indeed, one might think that the ability to decouple from the sensory flow is a later achievement, both phylogenetically and ontogenetically (Pezzulo, 2011; Pezzulo and Castelfranchi, 2009). In contrast to the idea that imagination is basic, perception would, in this view, emerge long before imagination, and it is the latter that reuses the resources used for the former and not the other way around.

To put this another way, while our experience at any given time is determined by the multilevel predictive model that our brains have selected, sometimes this is in order to make sense of imminent sensory input, and what we are talking about in this case is *online* cognition, which gives rise to *perceptual* experience. At other times, the story goes, the link to the outside world is broken. This would be achieved through a down-regulation of precision, and what we are talking about here is *offline* cognition. The natural thing to say, one might argue, is that one subspecies of such offline cognition is *imaginative* experience. Other subspecies might include, for

example, dreaming (Hobson and Friston, 2012), mind-wandering (Metzinger, 2018), hallucination (Wilkinson, 2014), or inner speech (Wilkinson and Fernyhough, 2017). Some might construe these as forms of imagination, or as involving imagination, others not.[3] In what follows we will focus on a more restricted notion of deliberate imagination as a distinct form of offline cognition.[4]

This overall approach views imagination proper, not so much as an all-pervasive building block of experience, but as a particular subspecies of offline cognition, that, yes, has a lot in common with, say, perception, but is to be sharply distinguished from it in a number of ways. The main challenge for accounts of imagination then becomes what the precise nature of that distinction is.

But even what perception and imagination have in common is qualitatively different, and this reflects a real virtue of the PPF in explaining a feature of imagination (and related mental episodes). Because the hypothesis-building in the imaginative (decoupled) instance is not driven by an attempt to minimize incoming prediction error, it is not hypothesis-building at the very bottom of the hierarchy (namely, at or near the organism's sensory surfaces), and hence the corresponding experience has certain phenomenological features. For example, in imagining a blue car, you might not imagine necessarily, as you do in perception, that particular shade of blue. In the broadest terms, imaginative experience tends to be less vivid and determinate, and more abstract, and the PPF has a convincing explanation of these differences.[5]

Departing from Reality Without Surprise

An immediate apparent worry arises for the PPF account of imagination. According to the PPF, our experiential content arises from the interaction between top-down predictions and error signals from our sensory contact with *reality*. One of the most obvious problems for accommodating imagination within the PPF is that the purpose of imagination is, in most cases, to envisage a *departure* from reality. This raises the worry of how we are able to *maintain* imagined content despite its conflict with the incoming sensory signal. One would expect predictions of content that departs from reality to generate significant error signals and thus be quickly extinguished as the relevant error signals propagate through the hierarchy. After all, the predictive hierarchy works to minimize overall surprise and imaginative departures from reality are often, in the technical sense, surprising. Thus, the question arises as to how imaginative content can be maintained despite conflicting with our sensory input. If our minds are made to model and track reality, how do we accomplish the sustained suspension of disbelief? It is hard to explain how we can foster a state of what Keats called "negative capability . . . when man is capable of being in uncertainties, mysteries, doubts, without an irritable reaching after fact" (Keats, 1958: 193–194).

---

3 For example, Wilkinson and Fernyhough (2017) have argued that it is misleading to think of inner speech as involving imagination.

4 We will henceforth refer to this simply as "imagination," but this does not indicate that we see this as the only mental process worthy of the name.

5 Imagination doesn't *have* to be less vivid or determinate than perception, as we can conceive of super-imaginers who can imagine with perceptual vivacity. However, this would be an exceptional case.

This worry can potentially be assuaged by taking account of the role of precision-weighting according to the PPF. Earlier we saw how turning down the precision on incoming error signals was central to explaining action in PPF. However, it may also play a central role in explaining how we are able to maintain imaginative content that is mismatched with the immediate environment. When we are engaged in an imaginative exercise, we should expect there to be a large discrepancy between the imagined content being predicted by higher levels in the hierarchy and incoming sensory signals. As such, we expect the incoming sensory signals to be extremely imprecise with respect to our *departure* from reality. Therefore, by assigning a low precision-weighting to error-signals coming from the sensory periphery, the errors can be down-modulated to the extent that they do not perturb the imagined content. Just as in action control, where precision-weighting allows one to maintain content of an action that has not yet happened in order to bring that action about, in the case of imagination, precision-weighting allows one to maintain content pertaining to circumstances that are not (yet) occurring in one's immediate environment and that may never actually occur.

However, while this explains how imagination can be *maintained*, it does not explain how it can be *constrained*. Our imaginative episodes tend to be relatively coherent, so one would expect there to be something playing the error-correcting role of inputs from the external world and thereby constraining the *development* of imaginative activity.

Imaginative Agency and Imaginative Constraints

An upshot of this appeal to precision-weighting is that imagination is just like perception but without being constrained by incoming sensory input. However, this immediately gives rise to the question of what, if anything, *constrains* imagination. The deliberate act of imagining can be distinguished from dreaming, hallucination, and mind-wandering by the fact that it is a product of our personal-level agency. Imagination is often *voluntary* and *goal-directed*, and must therefore be *constrained*.

Imaginative Agency

Imagining is something we often do voluntarily, usually to serve a particular purpose. One way of appreciating this phenomenon is to carefully contrast imagination with imagery. In contrast to imagery, imagination is a personal-level phenomenon: People are engaged in acts of imagination, people imagine things, and, although they do all manner of things that *involve* imagery, they don't do imagery, or engage in imagery *tout court*. Acts of imagination enable people to appreciate, in potentially many different ways, non-actual scenarios, and, when they are engaged in such acts, they may be motivated to do so by a number of different things. These acts of imagination will often recruit or make use of imagery in many modalities, but there

will also be aspects to the imaginative experience that aren't imagistic (Langland-Hassan, 2015; Yablo, 1993).[6]

Imagination is often within the control of our intentions; it is something that we do for a reason (Langland-Hassan, 2016). We use our imagination to serve a variety of different purposes. We may be trying to remember the color of someone's hair, judge whether we could jump over a stream, reason about a social situation, or simply entertain fantastical scenarios for the pleasure of it. The potential aims of imagination include (but are not limited to) (1) knowledge-acquisition, (2) creativity, and (3) comprehension of others' creative outputs. In short, imagination is goal-directed.

Imaginative Constraints

Understanding imagination as a goal-directed activity immediately gives rise to issues regarding how we are able to keep the imagination under control so that it reaches its goals. If imagination were just the free association of ideas, then it's hard to see how we could ever plot a stable course to the goal that we are aiming for. Imagination must be constrained in some way so that we can pursue a given goal. However, all of the goals that we put imagination to require that it is in some sense *generative*. To serve any useful purpose it must give us more than we already started with. Thus, in order for imagination to serve its purpose, it cannot be too constrained by our intentions. If we could only imagine exactly what we intend to imagine, then imagination couldn't give us any new content (Langland-Hassan, 2016). To explain the balance between the goal-directed and generative nature of imagination, the PPF needs to explain how imagination is constrained to just the right extent.

This is particularly pertinent when one focuses on cases in which the imagination is used to generate knowledge. The imagination often *seems* unbounded and without constraint. However, if this were the case then it would be *more* likely to generate falsehoods than truths. Imagination must be constrained in order to provide knowledge (Balcerak Jackson, 2018; Kind, 2016; Williamson, 2016). On the other hand, if we want to generate any new knowledge, what we imagine cannot be entirely "up to us," as it must go beyond what we already know (Balcerak Jackson, 2016, 2018; Langland-Hassan, 2016). The knowledge-generating capacity of the imagination can be explained by suggesting that some constraints on the imagination are fixed and within our control, allowing for exploratory activity within these constraints. For example, if one wants to know whether a sofa will fit through a door, one must constrain one's imagination to keep the relevant shapes and sizes and the laws of physics fixed, while allowing for exploration of various possible ways of manipulating the sofa. Van Leeuwen refers to this kind of imaginative process as "*exploratory constraint satisfaction*" (van Leeuwen, 2013: 229).

Imagination does not only serve to provide us with knowledge, it also plays a central role in creative activity, for example in the development of art and fiction for others' consumption. As with the case of knowledge-directed imagination,

---

6 It should be noted that some take imagination to be inherently imagistic (Kind, 2001); however, this does not preclude a distinction between imagination and imagery, as there can be imagery that is not the product of imagination.

creative imagination must be constrained. The structures, situations, or narratives that play out in our minds are not wild free-associations of concepts, they are coherent, in some sense that is akin to the coherence of perceptual experience, albeit with room for departures from reality. Again, the notion of imaginative exploration within chosen constraints seems apt to capture this activity. Creative imaginers explore what is possible within well-defined constraints that behave in accordance with certain (natural or narrative) expectations.

A further role for imagination lies in our ability to engage with others' creative outputs. When we engage with a fiction, our imaginative episodes are, again, constrained in various ways. First, our imaginative episodes tend to try to conform to our beliefs about reality as much as possible (Weisberg and Goodstein, 2009), although not necessarily the immediately experienced environment. Second, imaginative engagement with fiction is constrained by expectations determined by genre-specific narrative conventions (van Leeuwen, 2013).

Where are the Constraints in PPF?

The problem for the PPF account of imagination lies in explaining where the voluntary constraints on imagination come from. How can the imagination be constrained to achieve its goals without being kept in check by the outside world, when, according to the PPF, the constraints on perceptual inference come from error signals from the world? Both perception and action are dynamic processes, with experiential content unfolding over time in a relatively coherent manner. In the case of the former, this coherence can be explained by the fact that the content is constantly being entrained to reality by an incoming error signal. However, in the case of imagination, where the precise aim is a departure from reality, there is no error signal coming from a merely possible world to keep the content in check. In line with this concern, even Clark admits that "it seems very likely that for most creatures acts of deliberate imagining . . . are simply impossible" (Clark, 2016: 94). Yet acts of deliberate imagining are precisely the target explanandum.

This problem is not limited to accounts of the imagination. It is equally problematic in accounting for any departure from one's immediate reality. The PPF should be able to account for how we are able to accurately remember the past despite the sensory signals having long departed, and how we are able to predict and plan for distant future events taking place in environments different than those from which one's current sensory stimulation is coming.[7] In these cases, as with imagination, the question arises as to where the relevant error signals for constraining one's departure from immediate reality come from.

The PPF can arguably explain these capacities for "mental time travel" in terms of dampening the error signals from sensory inputs by lowering the expected sensory precision, so that error signals from *within* the generative hierarchy, rather than the sensory periphery, play the primary role in constraining content (see Clark, 2016:

---

7 There is some evidence to suggest that memory, planning, and imagination may have more in common than is often assumed, suggesting that all of these processes that call for departures from reality may be mediated by the same underlying mechanisms (De Brigard, 2014, 2017).

102–108). In this way, content can be constrained by past occurrences, since the hierarchy has been shaped by experiences, and more distant future predictions can be kept in check by the inbuilt assumption that the future will resemble the past that has sculpted the hierarchy.

Returning to the case of imagination, for situations that are similar to those that we have encountered in the past, imagination may similarly be constrained by error signals from within the hierarchy in the same manner as in cases of memory or longer-term prediction. Intra-hierarchical error signals may explain our ability to imagine scenarios that are closely related to our immediate situation or those that closely mirror our own past experiences. Imagination is thus likely to work in a similar fashion to memory or planning when the goals of imagination are closely aligned with those of these other activities. However, it seems unable to account for our ability to purposefully engage with imaginative content that drastically departs from reality, either for gaining knowledge of more distant possibilities or for engaging in more fantastical creative or recreational imaginative episodes.

Imagination as Mental Action

So far, we have considered accounts that conceive of imaginative experiences as grounded in sensory hypotheses. Although this speaks to the sensory phenomenology present in much (although arguably not all) imagining, it generates problems with accounting for how these hypotheses can be purposeful and constrained. However, an alternative way of understanding imagination within the PPF is in terms not so much of perception as of action.

Active Inference and Mental Action

Imagination, in this view, emerges as predicted consequences of anticipated possible actions. In order to explain how actions are selected from a range of options, proponents of the PPF arguably need to include *counterfactual* representations (Burr and Jones, 2016; Friston et al., 2012; Seth, 2014), i.e. predictions of what the sensory consequences *would* be if one *were* to engage in a particular action. For example, in order to decide whether to look left or right, one must predict the likely sensory consequences of each option. Since many actions will be mutually exclusive, many such representations will inevitably be about merely fictional circumstances, representing possible sensory consequences of actions that never occur. The fact that action may already require representing non-actual states may provide a clue as to the basis of imagination, for which representation of merely possible or non-actual states is central (Pezzulo, 2011; Pezzulo and Castelfranchi, 2009). Imagination may emerge from active inference with "suppression of overt sensory and motor processes" (Pezzulo, 2012: 1; see also Pezzulo, 2017).

This reliance on action may seem like a perplexing move, since action is usually *contrasted* with sensory, perception-like events. Thus, at this point, it is vital to distinguish between two different senses of "action." In one sense (most commonly

used in the cognitive sciences), "action" simply means motor activity, the endogenous bringing about of bodily movement. But this is neither sufficient, nor is it necessary, for "action" in the second sense (most commonly used in philosophy), which is tied not so much to the *motoric,* but to the *agentive.* Roughly, something is an action if it is intentional (or, to be more precise, has a description under which it is intentional). If you deliberately imagine a red triangle (maybe you've been offered a large sum of money to do so), that counts as an action in the agentive sense, but not (at least not obviously) in the motoric sense.[8]

What is the relationship between these two senses of "action" within the PPF? We said in the section "From Perception to Action" (above) that motor activity is understood within the PPF in terms of *active inference* (where this is contrasted with perceptual inference). In active inference, the world (in this case the body) is brought in line with predictions (in this case proprioceptive predictions), instead of the other way around, as it is with perceptual inference. However, if precision is turned down, the prediction doesn't have to be fulfilled, the bodily movement doesn't ensue, and you end up with motor imagination or rehearsal. But although active inference is usually thought of in the context of explaining motor activity via proprioceptive predictions, there is nothing special or different about proprioception, and the same can apply to the sensory domain. In other words, the organism can generate sensory predictions that also don't need to be fulfilled, and the result here would be sensory imagination.

Note that, although this looks like a view of imagination as merely decoupled perception, it is grounded in active inference, not in perceptual inference. In relying on active inference, these events, both proprioceptive and sensory, have to be potential events. More specifically, they are decoupled versions of potential actions (in the philosophical sense), both practical and epistemic. And this provides constraints that you wouldn't get from freewheeling perceptual inference: Proprioceptive and sensory imagination in this view is constrained by potential bodily actions, by potential acts of listening and looking. Note that, in being active, visual imagination is best thought of as decoupled looking, rather than seeing; and auditory imagination is best thought of as decoupled listening, rather than hearing. We do not get our attention passively grabbed by what we imagine (as we do with a worldly event like a bang or a flash). We bring it actively (and often effortfully) into being. It is driven by our agency.

This core idea, that imagination involves possible actions and experiences, generates constraints that come from two main sources. The first is bodily constraints that are the result of the organism's phenotype. The second is constraints from past experience. Some of the ways in which actions affect sensory input are relatively stable, both throughout one's lifetime and over recent evolutionary history, since they are determined by the shape of one's body. These stable relationships between action and sensation are known as sensorimotor contingencies (O'Regan and Noë, 2001; Seth, 2014). Given this stability, they are likely also to be stable features of one's predictive model of the world (Burr and Jones, 2016). As a result, when one

---

8 That is why the latter is not always necessary for the former, and indeed, there is a great deal of action (motor activity) that is not action (intentional/agentive).

engages in merely imagined action, the content of one's imagination will be constrained by the sensorimotor contingencies for the kind of body that one possesses.

However, some constraints on the predicted sensory consequences of action derive from stable structure in the world rather than one's own body. For example, by engaging in the action of eating an apple on several occasions, we may come to associate this action with a certain gustatory consequence: the taste of an apple. These kinds of stable relationships can also be modeled by the predictive hierarchy, allowing the predicted consequences of actions to be constrained by the way the world is. As a result, the merely simulated action involved in imagination can also be suitably constrained by past experience.

Distant Fantasies and the Role of Language

The active-inference-based account improves upon the purely perceptual-inference-based account of imagination by accommodating the purposive nature of imagination, as well as the idea that imagination can be constrained. By deliberately simulating actions that never actually take place we can generate content that departs from our immediate circumstances, yet which is constrained by the brain's model of how actions impact on sensation.

However, one concern goes, in doing so, does this not still tie imagination too closely to the real world? The active-inference-based accounts can explain how we are able to simulate the kinds of actions that we tend to engage in with the kinds of bodies that we possess and the kinds of lives we have led, and this may account for a wide range of the kinds of imaginative activity that we engage with in everyday circumstances. However, even a cursory reflection on the phenomenology of imagination suggests that not all of our imaginative episodes are quite this mundane.

Thus far, neither a perceptual-inference-based nor an active-inference-based PPF account seems apt to explain the full range of our imaginative competences. Neither approach seems able to walk the tightrope between imagination being entirely unconstrained and freewheeling, on the one hand, and too closely tied to reality, on the other. A potential solution may lie in exploring the role that language plays in structuring our imaginative episodes.

Although Clark is skeptical about the possibility of deliberate imagination in most creatures, he takes deliberate imagination to be possible in humans as a result of "the use of self-cueing via language" (Clark, 2016: 94). In many cases, our ability to generate imaginative episodes is closely tied to our consumption and production of *narrative*. Perhaps language helps to steer attention when we engage in offline imagination, much as it can do in our online interactions with the world (Lupyan, 2012; Lupyan and Clark, 2015).

As humans, we develop in a linguistically saturated environment. Language serves as more than just a medium for communicating thought; it is also central to the development of our cognitive capacities and to shaping and steering our thoughts (Clark, 2006). The hierarchical generative model that we build up through our interactions with the environment models more than just the natural causal structure of the world around us: It is also likely to encompass associations between words and

their contents. Simply hearing, reading, or merely entertaining a word or sentence about a particular content is likely to impact the predictions regarding perceiving or acting on instances of that content. Words create "artificial contexts" in which one is biased toward expecting to perceptually encounter their content (Lupyan and Clark, 2015: 282–283). For example, merely hearing the word that corresponds to a stimulus that would otherwise remain invisible to subjects can boost the stimulus into visual awareness (Lupyan and Ward, 2013).

In more mundane imaginative exercises, such as imagining whether a sofa will fit through the door, we tend not to engage in drastic fantastical departures from reality, and that is a very good thing. It wouldn't help if one suddenly imagined the sofa turning into a hippopotamus. Notice, however, that as you read the previous sentence, it was hard not to imagine the aforementioned fantastical transformation, despite it being far from our usual bodily interactions with the world. This suggests that the interaction between language and imagination may provide a route for the PPF proponent to account for our ability to imagine more distant fantasies. Linguistic cues may bias our offline simulations in a similar way to how they bias perception, bringing particular content to awareness. Furthermore, the compositional and thus productive nature of language may explain how we are able to combine contents in imagination of which we have no experience and that we are vanishingly unlikely ever to experience.

The idea that language has a significant role to play in a PPF account of imagination seems promising. However, much more work needs to be done to fully flesh out this idea. At first sight, the proposal seems to suffer from a "chicken-and-egg" problem. We tend to see narrative linguistic expression as the product of creative imaginative endeavors, so it is hard to see how language could be both the product of the process and the precursor that needs to already be in place to guide and constrain it. Moreover, the precise details of how complex linguistic abilities are realized according to PPF and how such abilities impact on other mental processes within the PPF remain to be worked out in detail. Thus, exploring the way in which language engenders the capacity for drastic imaginative departures from reality is a fruitful line for further research.

Concluding Remarks

On the one hand, PPF has the potential to demystify the imagination, as it frames the imagination as much more similar and closely tied to perception and/or action. On the other hand, a PPF account of the imagination faces problems of its own. In particular, as things stand, PPF must do more to account for how the imagination can be both suitably constrained and sufficiently generative without the relevant constraints being imposed by the world. How can our predictions be constrained by mere fictions and fantasies, when the main task of the brain is to predict reality rather than departures from it?[9] If the PPF can overcome these issues,

---

9 Answering this question could also provide insight into cases in which the mind is detrimentally constrained by fiction and fantasy in atypical contexts (e.g. delusions and hallucinations in schizophrenia). In particular, it might be that imaginative episodes in these contexts are, for whatever reason, not recognized as such by the subject.

perhaps by appealing to the role of language, and provide an account of the imagination, then it can come one step closer to fulfilling its promise of providing a grand, unified approach to the mind.

References

Balcerak Jackson, M. (2016). On Imagining, Supposing and Conceiving. In A. Kind and P. Kung (eds.), *Knowledge Through Imagination*. Oxford, UK: Oxford University Press, 42–60.

(2018). Justification by Imagination. In F. Macpherson and F. Dorsch (eds.), *Perceptual Imagination and Perceptual Memory*. Oxford, UK: Oxford University Press, 209–226.

Bayes, T. (1763). LII. An Essay Towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, FRS Communicated by Mr. Price, in a Letter to John Canton, AMFR S. *Philosophical Transactions of the Royal Society of London*, *53*, 370–418.

Burr, C., and Jones, M. (2016). The Body as Laboratory: Prediction-Error Minimization, Embodiment, and Representation. *Philosophical Psychology*, *29*(4), 586–600.

Clark, A. (2006). Language, Embodiment, and the Cognitive Niche. *Trends in Cognitive Sciences*, *10*(8), 370–374.

(2013). Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science. *Behavioral and Brain Sciences*, *36*(3), 181–204.

(2015). Perception as Prediction. In D. Stokes, M. Matthen, and S. Biggs (eds.), *Perception and its Modalities*. New York, NY: Oxford University Press, 23–43.

(2016). *Surfing Uncertainty: Prediction, Action and the Embodied Mind*. Oxford, UK: Oxford University Press.

Corlett, P. R., Taylor, J. R., Wang, X. J., Fletcher, P. C., and Krystal, J. H. (2010). Toward a Neurobiology of Delusions. *Progress in Neurobiology*, *92*(3), 345–369.

De Brigard, F. (2014). Is Memory for Remembering? Recollection as a Form of Episodic Hypothetical Thinking. *Synthese*, *191*(2), 155–185.

(2017). Memory and Imagination. In A. Kind (ed.), *The Routledge Handbook of Philosophy of Imagination*. London, UK; New York, NY: Routledge, 127–140.

Feldman, H., and Friston, K. (2010). Attention, Uncertainty, and Free-Energy. *Frontiers in Human Neuroscience*, *4*, 215.

Friston, K., Adams, R., Perrinet, L., and Breakspear, M. (2012). Perceptions as Hypotheses: Saccades as Experiments. *Frontiers in Psychology*, *3*, 151.

Hobson, J. A., and Friston, K. J. (2012). Waking and Dreaming Consciousness: Neurobiological and Functional Considerations. *Progress in Neurobiology*, *98*(1), 82–98.

Hohwy, J. (2013). *The Predictive Mind*. Oxford, UK: Oxford University Press.

Keats, J. (1958). *The Letters of John Keats, 1814–1821*. Edited by H. E. Rollins. Cambridge, MA: Harvard University Press.

Kind, A. (2001). Putting the Image Back in Imagination. *Philosophy and Phenomenological Research*, *62*(1), 85–109.

(2016). Imagining Under Constraints. In A. Kind and P. Kung (eds.), *Knowledge Through Imagination*. Oxford, UK: Oxford University Press, 145–159.

Kirchhoff, M. D. (2018). Predictive Processing, Perceiving and Imagining: Is to Perceive to Imagine, or Something Close to it? *Philosophical Studies*, *175*(3), 751–767.

Langland-Hassan, P. (2015). Imaginative Attitudes. *Philosophy and Phenomenological Research*, *90*(3), 664–686.

(2016). On Choosing What to Imagine. In A. Kind and P. Kung (eds.), *Knowledge Through Imagination*. Oxford, UK: Oxford University Press, 61–84.

Lupyan, G. (2012). Language Augmented Prediction. *Frontiers in Psychology*, *3*, 422.

Lupyan, G., and Clark, A. (2015). Words and the World: Predictive Coding and the Language-Perception-Cognition Interface. *Current Directions in Psychological Science*, *24*(4), 279–284.

Lupyan, G., and Ward, E. J. (2013). Language Can Boost Otherwise Unseen Objects into Visual Awareness. *Proceedings of the National Academy of Sciences*, *110*(35), 14196–14201.

Metzinger, T. (2017). The Problem of Mental Action: Predictive Control Without Sensory Sheets. In T. Metzinger and W. Wiese (eds.), *Philosophy and Predictive Processing*. Mainz, Germany: Johannes Gutenberg-Universität.

(2018). Why is Mind-Wandering Interesting for Philosophers? In K. Christoff and K. C. R. Fox (eds.), *The Oxford Handbook of Spontaneous Thought: Mind-Wandering, Creativity, and Dreaming*. Oxford, UK: Oxford University Press, 97–111.

O'Regan, J. K., and Noë, A. (2001). A Sensorimotor Account of Vision and Visual Consciousness. *Behavioral and Brain Sciences*, *24*(5), 939–973.

Pezzulo, G. (2011). Grounding Procedural and Declarative Knowledge in Sensorimotor Anticipation. *Mind & Language*, *26*(1), 78–114.

(2012). An Active Inference View of Cognitive Control. *Frontiers in Psychology*, *3*, 478.

(2014). Why Do You Fear the Bogeyman? An Embodied Predictive Coding Model of Perceptual Inference. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(3), 902–911.

(2017). Tracing the Roots of Cognition in Predictive Processing. In T. Metzinger and W. Wiese (eds.), *Philosophy and Predictive Processing*. Mainz, Germany: Johannes Gutenberg-Universität.

Pezzulo, G., and Castelfranchi, C. (2009). Thinking as the Control of Imagination: A Conceptual Framework for Goal-Directed Systems. *Psychological Research PRPF*, *73*(4), 559–577.

Seth, A. K. (2014). A Predictive Processing Theory of Sensorimotor Contingencies: Explaining the Puzzle of Perceptual Presence and its Absence in Synesthesia. *Cognitive Neuroscience*, *5*(2), 97–118.

van Leeuwen, N. (2013). The Meanings of "Imagine" Part I: Constructive Imagination. *Philosophy Compass*, *8*(3), 220–230.

Weisberg, D. S., and Goodstein, J. (2009). What Belongs in a Fictional World? *Journal of Cognition and Culture*, *9*(1–2), 69–78.

Wilkinson, S. (2014). Accounting for the Phenomenology and Varieties of Auditory Verbal Hallucination Within a Predictive Processing Framework. *Consciousness and Cognition*, *30*, 142–155.

Wilkinson, S., and Fernyhough, C. (2017). Auditory Verbal Hallucinations and Inner Speech: A Predictive Processing Perspective. In Z. Radman (ed.), *Before Consciousness: In Search of the Fundamentals of the Mind*. Exeter, UK: Imprint Academic, Limited, 372–397.

Williamson, T. (2016). Knowing by Imagining. In A. Kind and P. Kung (eds.), *Knowledge Through Imagination*. Oxford, UK: Oxford University Press, 113–123.

Yablo, S. (1993). Is Conceivability a Guide to Possibility? *Philosophy and Phenomenological Research*, *53*(1), 1–42.