Philos Stud (2018) 175:751–767
https://doi.org/10.1007/s11098-017-0891-8

Predictive processing, perceiving and imagining: Is to perceive to imagine, or something close to it?

Michael D. Kirchhoff[1]

Published online: 10 March 2017

Abstract This paper examines the relationship between perceiving and imagining on the basis of predictive processing models in neuroscience. Contrary to the received view in philosophy of mind, which holds that perceiving and imagining are essentially distinct, these models depict perceiving and imagining as deeply unified and overlapping. It is argued that there are two mutually exclusive implications of taking perception and imagination to be fundamentally unified. The view defended is what I dub the ecological–enactive view given that it does not succumb to internalism about the mind-world relation, and allows one to keep a version of the received view in play.

Keywords Predictive processing · Imagination · Perception · Internalism · Embodiment · Enactivism · Realization

1 Introduction

Many philosophers think that there is a sharp distinction between imagination and perception. No doubt this is due to the following: to imagine X is to think of X, where X is not served up by the world in perception. Unlike perceiving, to imagine is to enact a mental activity at one remove from the real world. Many find this view, or something like it, convincing. If coherent, it grounds the intuition that imagining is nontrivially distinct from perceiving. Those compelled by this view often appeal to such an alleged difference to define imagination. Gendler (2011) sketches it as follows: "*To imagine* something is to form a particular sort of mental representation of that thing. Imagining is typically distinguished from mental states such as

✉ Michael D. Kirchhoff
kirchhof@uow.edu.au

1 Department of Philosophy, University of Wollongong, Wollongong, Australia

perceiving, *remembering* and *believing* in that imagining *S* does not require (that the subject consider) *S* to be or have been the case, whereas the contrasting states do." (2011, p. 1) Or, as Stevenson says: to imagine is "the ability to think of something not presently perceived, but spatio-temporally real" (2003, p. 238). Given that many contemporary philosophers accept this view about perceiving and imagining, I will call the view that imagining is essentially distinct from perceiving 'the received view'.

A new theory in neuroscience and philosophy of mind threaten to dampen the initial appeal of the received view, because it "suggests… a deep unity between perceiving and imagining" (Clark 2014, p. 39). This theory states that the brain is a homeostatic prediction-testing machine (Clark 2012; Clark 2013a, 2014, 2015; Friston 2005, 2009, 2010, 2013; Friston and Stephan 2007; Friston et al. 2012a; Hohwy 2013, 2014, 2015). It provides a model of the brain as a proactive and self-organizing system: one whose function is to reduce discrepancies between the ways it predicts the world to be, and the way the world (i.e., sensory input) actually is. As Colombo and Wright point out, the "mechanistic function of minimizing prediction error is constituted by various monitoring- and manipulation-operations on hierarchical, dynamical models of the causal structure of the world within a bidirectional cascade of cortical [and sub-cortical] processing." (2016, p. 1) Call this view 'Predictive Processing' (PP, for short). PP may be said to represent a departure from the received view by depicting perceiving and imagining as deeply unified and overlapping.[1]

Defenders of PP are excited about the potential implications of this theory. In Friston's words, "if one looks at the brain as implementing this scheme, nearly every aspect of its anatomy and physiology starts to make sense." (2009, p. 293). Or, as Clark says about PP, it provides "a deeply unified theory of perception, cognition, and action." (2013a, p. 186) Hohwy is equally optimistic, as he says: it explains "everything about the mind." (2015, p. 1)[2] I do not pretend to cover the potential of PP being a grand unifying theory. Moreover, the discussion will be cast at a relatively abstract level, examining particular philosophical implications of conceiving of perception and imagination as continuous and deeply unified. Let us call the idea that perception and imagination are fundamentally overlapping and unified the 'deep unity theory' (DUT, hereafter). In what follows, I argue that there are two mutually exclusive readings of the implications of the DUT. The first I call the *inferred fantasies view*. I call the second the *ecological–enactive view*. I briefly set out both here.

The first articulation of DUT integrates perception with imagination, and vice versa, given (1) the idea that perception and imagination are the psychological results of the brain endogenously generating its own sensory inputs top-down, and

[1] Commenting on the general idea of unity and continuity, Seth states: "More generally, the key role of top-down predictive or generative models [predictive processing] in perception points to a strong continuity not only with imagery but also with associative recall, dreaming, and other self-generated perceptual or quasi-perceptual states" (2014, p. 101).

[2] See Colombo and Wright (2016) for an excellent critical discussion of PP and its ambitions of being a grand unifying theory.

(2) the view that perceiving is understood as an "inferred fantasy about what lies behind the veil of input" (Paton et al. 2013, p. 222), or as a process of "controlled hallucination" (Clark 2013a, p. 25). This is the inferred fantasies view. Now, this is not an attempt to argue that perception and imagination are identical. Instead, the inferred fantasies view puts forth the idea that perceiving is akin to imagining precisely because perceiving is conceived of as "like a fantasy or virtual reality constructed to keep the sensory input at bay." (Hohwy 2013, p. 137) For some defenders of this view, it implies a commitment to internalism about the mind-world relation. I restrict the claim that the inferred fantasies view implies internalism of a vehicular kind. Nowhere is this more evident that in the following quote by Hohwy: "PEM [prediction error minimization] reveals the mind to be inferentially secluded from the world … to be … skull-bound, and "open the door to skepticism" (2014, p. 1). Note that I do not treat this claim as an objection to the inferred fantasies view. Vehicle internalism is a substantial philosophical position that cannot be used as an argument in and of itself against a specific theory or perspective. What I shall argue is that there is an articulation of DUT available that need not lead to internalism about the mind-world relationship: this is the ecological–enactive view.

The ecological–enactive view suggests (pace Clark 2014) that there is a fundamental linkage between perception and imagination. On this formulation, perception and imagination are distinct phenomena, yet unified aspects of the same underlying machinery for prediction error minimization. This is a shared assumption with the inferred fantasies view. However, it does not imply internalism with respect to the mind-world relation. To see this, consider, as a first pass, any chair in your house. The chair is something that exists—it is right there, in your living room. When you perceive the chair, you experience the chair as real, "as existing in the world." (Seth 2014, p. 97) You also experience the chair as affording possible actions. That is, how the world is and your capacities for action really matter to perception (Noë 2004). Such features of perception are not well accounted for through the lens of vehicle internalism (Ward 2012). Indeed, here perception is depicted as constrained—in a way that imagination is likely not to be—by an organism's embodiment and the environmental context. Disentangling perceiving and imagining in this way yields a way to ward off the claim (or so I shall argue) that DUT implies internalism about the mind-world relation, and to keep a version of the received view in play.

The rest of the paper has four sections. In Sect. 2, I give a brief presentation of PP. I do so briefly, since I will elaborate PP in further detail as I move through the paper. In Sect. 3, I discuss the inferred fantasies view. In Sect. 4, I develop the ecological–enactive view. I argue that there are reasons to prefer the latter at the expense of the former. I conclude in Sect. 5.

2 Predictive processing: the view, in brief

PP models depict the brain as a prediction machine (Clark 2016; Hohwy 2014; Friston 2010). In this discussion, prediction is used explicitly as the weighted mean of a random variable from probability theory and statistics. Here, a weighted mean of a random variable is a “magnitude posited to be transmitted downwards as a driving signal by the neurons comprising pair-wise levels in the cortical hierarchy.” (Colombo and Wright 2016, p. 2) It is in this sense that brains are understood to be organs of prediction. So conceived, the brain is understood as constantly seeking to minimize a prediction error quantity, while, at the same time, trying to estimate and make more precise the reliability of the prediction error signal (Hohwy 2012). Here the term ‘prediction error’ is a measure of discrepancy between top-down predictions about the value of some random variable and its actual value. In other words, it refers to a mismatch between some specific prediction about sensory input and the actual input received (Clark 2015). Precision is effectively a measure of uncertainty. In the context of prediction error minimization, precision estimation is what adjusts the amplitude (or volume) of prediction errors. Indeed, “a more precise prediction error will have more weight on updating the system’s models of the world.” (Colombo and Wright 2016, p. 2) Under the prediction error minimization scheme, agents can be cast as always seeking to optimize the conditions for avoiding entering into states with high surprise. Surprise is defined as a measure of improbability (Friston et al. 2012b) in the sense that it refers to the negative log probability of an outcome (Friston 2009). Surprise is the long-term average of prediction error. So, by minimizing prediction error, agents are equally optimizing the predictability of the states that they will find themselves in over time. In this context, surprise is measured relative to a generative model. A generative model is a statistical model of how sensory input are generated, which is constituted by two components: (1) prior probability distributions over bodily and/or environmental states; and (2) the likelihoods of an agent’s sensory input given its bodily and/or environmental causes (Colombo and Wright 2016). In sum, PP is the view that psychological phenomena come about through the same process: minimization of prediction error.

The brain is said to realize prediction error minimization in virtue of functioning in accordance with approximate Bayesian inference (Friston 2010; Hohwy 2015). To explain this, I refer to a simple example.

Suppose you walk into a hardware store and ask: could I get *four candles*? Were the shopkeeper to hand you *fork handles*, you would be surprised. We can explain why she hears you saying fork handles even though you said four candles by appeal to Bayesian inference. Assume she sells many more fork handles than she does candles. If so, we may assume that she does not even hear you saying ‘four candles’ but hears you saying ‘fork handles’ (Stone 2013). How could this be so? In this example, the evidence (the acoustic data) is ambiguous. It is consistent with two different predictions as a result of the prior probability distributions and the likelihood function of the agent’s statistical (generative) model given the data. Without knowing it, she (or better, her predictive brain) weighs one prediction higher than the other. Doing so assigns a higher *posterior probability* to one of the two predictions: the probability of you saying fork handles given the acoustic data. If this is true, her subpersonal predictions align with Bayes’ theorem.

In formalizing probabilities, one must start by introducing a hypothesis space. A hypothesis space is a probability distribution. A probability distribution comprises different hypotheses (or predictions), and a probability function is assigned to each hypothesis, p(h). This reflects the probability of h (Rescorla 2013). The conditional probability of each hypothesis is: the probability of the acoustic data given that 'four candles' was uttered, and the probability of the acoustic data given that 'fork handles' was uttered. In Bayesian terms, this is known as *likelihood*, which is a function of the parameters of a generative model given the data. Yet, likelihood is not sufficient to generate the posterior probability. We want to know the posterior probability. After all, we already know that the shopkeeper has heard the acoustic data. Hence, we really want to know the *inverse* of the conditional probability: the probability of the hypothesis given the acoustic data. If we know this, we know the posterior probability of the hypothesis in question, where a posterior probability is an updated probability measure of h (Rescorla 2013). With this in mind, we can say, following Stone, that without "knowing it, [the shopkeeper] probably used … Bayes to her what the customer probably said." (2013, p. 17).

Indeed, the shopkeeper hears you saying 'fork handles' because her brain meets the sensory input with a strong prediction about 'fork handles'. It is tempting to think of this case as one of misperception. Yet, this does not seem quite right when approached from the perspective of PP. It is simply what her brain expects about the world. That is, in this particular case, the probability of the 'fork handle' hypothesis is higher than any alternative hypothesis. This results in the 'fork handle' hypothesis being selected, thereby determining what she perceives. As Hohwy emphasizes: the "winning hypothesis about the world is the one with the highest posterior probability, that is, the hypothesis that best explains away [i.e., accounts for] the sensory input." (2014, p. 5)

On the basis of this overview of PP, I now turn to consider the implications of DUT in detail. I start by examining the inferred fantasies view.

3 The inferred fantasies view

According to what I dub the inferred fantasies view, perception and imagination are understood as deeply unified. There are at least two reasons for this. The first is that perceiving is akin to imagining because to perceive is to experience a fantasy or a virtual reality. The second is that perceiving, just like imagining, is epistemically secluded from the world. Both of these claims converge on a key assumption: that we do not enjoy direct access to the world around us. Perception, as with imagination, is indirect in the "sense that what we experience *now* is given in your top-down predictions of sensory input, rather than the bottom-up signal from the states of [worldly] affairs themselves." (Hohwy 2013, p. 48; italics in original) In a statistical sense, perception is indirect because it is effectively the result of the brain producing a statistical estimate, extracted from inferences over probability distributions comprising the brain's generative model (Hohwy 2013). In its most extreme form, the inferred fantasies view is argued to imply a form of Cartesian skepticism. As Hohwy states, because "we cannot obtain an independent view of our position in the world, we cannot exclude the skeptical hypothesis that the sensory input we receive is caused by an evil, hoaxing scientist [or the Bayesian brain] rather than the external states of affairs we normally believe in." (2014, p. 7)

One can argue for this view as follows. First, a signature aspect of the predictive brain is that it deploys generative models to minimize prediction error. As we saw above, a generative model is a statistical model of how sensory data are generated. It is comprised of two aspects: (1) prior probability distributions over bodily and/or environmental states; and (2) generative probability distributions (or likelihoods) of an agent's sensory input given its bodily and/or environmental causes (Colombo and Wright 2016). In the context of Bayesian inference, the probabilistic relation between priors and evidence (data) is usually articulated as "a prior belief and the likelihood of observations, given their causes." (Hobson and Friston 2014, p. 32) Second, prediction error minimization unfolds within a neural processing hierarchy. As Colombo and Wright put it: "this compact processing scheme leads to bidirectional processing, where feed-forward connections convey information about the difference between what was expected and what actually obtained—i.e., prediction error—while feedback connections convey predictions from higher processing stages to suppress prediction errors at lower levels. So, processing at each stage signals difference between expected and actual features to the next stage up the hierarchy; and each stage sends back to the one below it the expected features." (2016, p. 2) This process of hierarchical prediction error minimization comes to an end when the brain is able to meet the forward flowing error signals with a prediction that best explains away to the error signal itself (Hohwy et al. 2008).

This gives us the brain's putative job description. It must explain the sensory data it receives by generating top-down predictions that best match the input in question. The prediction that best explains the sensory input is also the prediction—at least in non-pathological cases—that has the least prediction error, which, in turn, is what determines perception, action, cognition, and so on. Now, prediction error can be evaluated as a function of two quantities that the brain has access to: prior probability distributions over bodily and/or environmental states; and generative distributions or likelihoods of an agent's sensory input given its bodily and/or environmental causes (Colombo and Wright 2016). Given that the brain does not have 'access' to the causes of its sensory input, the generative model functions as an approximate probability distribution of the causes of sensory input (Friston 2010). By utilizing generative models, brains are able to sample the approximate probability distribution—that is, the mean and standard deviation between the dependencies of causes and actual sensory input. It is this process that leads to prediction error. The brain's job, then, is to produce predictions that best explains away prediction error on the basis of its probabilistic generative models.

It is this process of utilizing an internally encoded generative model to lower prediction error that lead proponents of PP to the view that what we perceive is akin to a virtual reality. As Clark puts it: "the best way to anticipate/match the incoming signal is to discover and deploy internal resources that amount to a kind of '*virtual reality generator*' that models the distal elements and their typical modes of interaction." (Clark 2013b, p. 9; italics added)

This job description can be given a familiar Bayesian explanation. We perceive when the brain arrives at a hypothesis with the highest posterior probability. The hypothesis with the highest posterior probability outcompetes all other probable hypotheses—in the hypothesis space—and thus determines what we perceive.

Once this picture is in place, we can see how it leads to the claim that perceiving is unified with imagining in the sense that it engenders a view of perceiving as a virtual reality. For it is the brain's ability to self-generate its own data that enables perceiving. And its self-generated data just is an inferred fantasy that best (probabilistically speaking) explains the structures most likely to have caused the sensory signals it receives.

Another reason for thinking that the brain is a virtual reality generator has to do with *timing*. Since our brains must make predictions prior to accounting for prediction error, it is always one step ahead of its sensorium living forever (or so it would seem) in the immediate future. This is how Bar (2007) sees the issue, noting that rather than "passively 'waiting' to be activated by sensations, it is proposed that the human brain is continuously busy generating predictions that approximate the relevant future." (2007, p. 280) But, it is not merely about timing. I suspect that a different reason for the claim that perceiving is akin to a fantasy or virtual reality concerns *probability*. That is, the brain is able to settle on a specific hypothesis by sampling across its own self-generated probability distributions. When considering probabilities, it is standard to note that the probability of some event, say, is between 0 (i.e., it will not occur under any circumstances) and 1 (i.e., it is certain to occur under any circumstance). As a result, for any event, there is a necessary margin of error (or uncertainty) relative to any given prediction. So while timing is an important factor here, it is because of the uncertainty of any hypothesis selected, that self-generated predictions could only ever approximate the truth. It is for this reason, that probability, and not merely timing, is a substantial reason for the claim that the brain is a virtual reality generator.

If this account of the inferred fantasies view is on the right track, then two things follow. First, "to perceive … is to deploy internal resources capable of endogenously generating those same activation patterns [as imagining]…" (Clark 2014, p. 39). The logic is simple: if imagining deploys the same processes active during perceiving, then the underlying neural structures supporting perceiving should also support imagining (Kosslyn 1994). Second, if imagining is at one remove from the world, and if perceiving and imagining are unified, then perceiving (all things equal)—just like imagining—is "nevertheless at one remove from the real world" (Hohwy 2013, p. 137).

Although such an account aligns with PP, it also faces an immediate problem. If the brain is a virtual reality generator, then 'what we perceive' is unclear. There are two options. The first is that what we perceive is the structured and meaningful world itself *relative to the sensorimotor capacities of organisms* (Noë 2004). The other option is that what we perceive is the *virtually* structured and meaningful world generated by the brain's own top-down predictions. In general, if the outcome of the inferred fantasies view is that perception or perceptual experience is a virtual reality, then what perceivers perceive cannot be the world itself—the 'real' world which Hohwy alludes to in the quote above—but at best a world or reality created by the brain in virtue of it only having access to a generative model encoded by its internal states. This follows in virtue of brain's not having access to the world outside its own generative model. It is this assumption that motivates the claim that perception, in the same way as imagination, is at one remove from the real world.

Some remain skeptical about this second option, even if they endorse the basic idea that the brain is a virtual reality generator. They are skeptical of this option because it threatens to imply a separation of mind and world. Yet Hohwy (2014) embraces this view of the mind-world relationship. He says, on the basis of PP, that the "mind can then be understood in internalist, solipsistic terms, throwing away the body, the world and other people [epistemically, speaking]" (2014, p. 7). *Prima facie*, at least, this is one possible consequence of PP. For on this reading of PP, it is the inversion of generative models, and subsequent precision estimation, which determines what we perceive. Furthermore, and if this reading is correct, then it would seem to follow that generative models give rise to an epistemic curtain between perceiver, on the one hand, and the world beyond the generative models, on the other.

Clark is one of those who are skeptical about the putative implication inferred on the basis of PP by Hohwy about the mind-world relation. That is, Clark suggests (2013b; but see also Orlandi 2014) that insisting on a veil of experience—carving off mind from world—contradicts the phenomenology of perceptual experience. Should this proposal turn out to be correct, then we are back with the first option: that what we perceive is the structured and meaningful world itself (relative to the sensorimotor capacities of organisms), as opposed to the brain's best hypothesis. There are two horns to Clark's argument. The first is to acknowledge that "there is something right about the idea that our expectations are in some important sense the primary source of all the contents of our experience, even though such contents are constantly being checked, nuanced, and selected by the prediction error signals consequent upon the driving sensory input." (2013b, p. 199) The second is that, for Clark, there "is no sense, even assuming the prediction-driven account is accepted, in which *what* we perceive is the brain's best hypothesis. Instead, what we perceive is the world, as (hopefully) revealed by the best hypothesis." (2013b, p. 25; italics added)

I suspect that many would find this reply convincing. But this conviction would be premature for at least one reason. It denies what it presupposes. The reply assumes that the brain generates virtual realities. It also assumes that it is these steps of engendering virtual realities that determine what we perceive. Yet, at the same time, it denies that what we perceive is fantastical. I find this perplexing. The options available bifurcate. Either we are only privy to what the world impresses upon the inside of the experiential curtain given that what we perceive is the brain's self-generated virtual data of sensory causes. Hohwy (2014) defends this implication. Hence, this idea that the inferred fantasies yields both a version of indirect realism on the basis of PP, and an epistemically secluded view of mind from the rest of the body and world. Or, we are able to penetrate the experiential curtain, reaching through it, all the way out to the world (so to speak). Clark (2013b) defends this result. However, if one starts by accepting that the brain is a virtual reality generator (as Clark does), and that it the brain's endogenously generated predictions that determine the content of perceptual experience, then it is difficult to see how there could be room for the kind of 'open channel' view, which Clark wants to hang on to. In other words, it cannot *both* be that we perceive the world as revealed through sensory input and that what we experience is the upshot of probabilistic inferences over probability distributions. Of course, this is not a knockdown argument against Clark's position, but it does highlight (if correct) a potential tension with (1) accepting that the brain is a virtual reality generator, and (2) wanting to defend the idea that what we perceive is the actual world itself. Indeed, it seems to be an outcome of inferring or sampling over probability distributions that *what* we perceive just is the brain's best guess (so to speak), whether we are aware of it or not.

4 The ecological–enactive view

This brings us to the ecological–enactive view (EE view). Like the inferred fantasies view, it states that perception and imagination are unified and dual-aspects of a single strategy for prediction error minimization. To properly present this view requires a bit of work. However, this work will reveal an alternative to the internalist commitments of the inferred fantasies view, on the one hand, and a way to keep intact a version of the received view, on the other.

Rescorla (2013) sketches a common way of introducing Bayesian probability theory in which the brain is depicted as operating with uncertainty. Rescorla (2013) voices this view, as he says that Bayesian probability theory is a "mathematical framework that models reasoning and decision-making under uncertainty" (Rescorla 2013, p. 694). So, it is no great surprise that many find the latter a good reason for thinking that the brain operates in accordance with Bayesian probability theory.

Insofar as PP models align with Bayesian probability theory, they handle uncertainty in the sensory signal by appealing to prior probabilities of hypotheses. This view is initially very attractive. It provides an explanation for how the brain copes with the problem of underdetermination (Rescorla 2013). Hohwy explains the problem as follows:

> In our complex world, there is not a one–one relation between causes and effects, different causes can cause the same kind of effect, and the same cause can cause different kinds of effect… If the only constraint on the brain's causal inference is the immediate sensory input, then, from the point of view of the brain, any causal inference is as good as any other… The key issue is then that without any *additional constraints* the brain would not be able to perform reliable causal inference about its sensory input… It seems obvious [that to perform] causal inference about things… [the brain must draw] on a vast repertoire of *prior beliefs* (2013, pp. 13–14).

Despite its initial appeal, however, there is a problem with this view. First, given the clear focus on priors, PP models depict perceiving as driven almost completely from the top down on the basis of generative models (comprised of prior knowledge about the structure of the world itself). But, second, if the sensory barrage confronting the brain is truly underdetermined, then "the space of all possible hypotheses concerning the environment… is infinite." (Orlandi 2014, p. 90; see also Rescorla 2013) That is a problem.[3]

Why does the shopkeeper's brain entertain hypotheses about 'fork handles' and 'four candles', as opposed to some other hypothesis? Orlandi puts the question as follows: "So it is a question for Bayesian models why some initial hypotheses, out of all possible ones, are formed and entertained." (2014, p. 90) If the hypothesis space is infinite, then the nontrivial question to address is: how is the hypothesis space restricted? One possibility is to appeal to priors already occupying states in the overall hypothesis space. This appeal is almost certain to fail. For it is uncertainty of the hypothesis space, in the first place, that is the problem (Orlandi 2014). So it will not suffice to appeal to the hypothesis space. After all, if it is the unconstrained uncertainty concerning hypotheses occupying states in a probability distribution that is the problem, then one cannot limit the uncertainty of the probability distribution by appeal to the probability distribution itself. This suggestion, therefore, fails.

One promising suggestion is that the underdetermination problem is ultimately taken care of through hierarchical, empirical Bayes. It is precisely because the processing profile of PP is hierarchical rather than, say, combinatorial that the problem does not arise. In hierarchical systems, each processing level is continuously trying to predict the activity of the level below. As Hohwy describes the outcome and function of hierarchical predictive processing:

> The input to the system from the sense is conceived as prediction error and what cannot be predicted at one level is passed on to the next. In general, low levels of the hierarchy predict basic sensory attributes and causal regularities at very fast, millisecond, time scales, and more complex regularities, at increasingly slower time scales, are dealt with at higher levels … Prediction error is concurrently minimized across all levels of the hierarchy, and this unearths the states and parameters that represent the causal structure and depth of the world (2012, p. 2).

The emphasis on hierarchical, empirical Bayes is a step in the right direction. But an emphasis on the hierarchical profile of predictive brains does not entail that it is *only* this internal processing architecture that enables the brain to achieve a grip on the process of hypothesis selection. That is, the hierarchical nature of the brain may just as well be seen as substantially contributing to, without being the sole realizer of, dealing with the underdetermination problem. One proposal is that in addition to hierarchical generative models, one might also need to appeal to the driving sensory signal itself (Orlandi 2014, p. 92).[4]

[3] In fact, it is the familiar *frame problem* in artificial intelligence. It is not my intention to develop this issue in detail—a task for another occasion. Rather, I make use of the issue here to further motivate the ecological–enactive view.

[4] Orlandi's proposal is to ground PP in Natural Scene Statistics (NSS). She explains: "One of the fundamental ideas of NSS is to use statistical tools to study *not* what goes on inside the head, but rather what goes on outside—for example, what are the likely environmental causes of retinal images. NSS is interested both in what is more likely present in the environment or, more tractably, in how probable a given cause is, and in the relationship between what is in the world and the stimulus it produces" (2014, p. 63).

However, while this might be a step in the right direction, one worry is that it does not add anything new to the picture. As it turns out, the appeal to driving sensory signals—allowing for hypothesis selection—is already a core part of Bayesian inference. To see this, consider that even if driving sensory signals help to restrict the space of possible hypotheses, one is nevertheless left with having to appeal to priors in order to deal with the underdetermination problem. So it is not evident that one can fully motivate the EE view by appealing to the role of driving sensory signals rather than hierarchical generative models.

This suggests that something else is required to motivate the possibility of the EE view. One promising way to go is to reconsider the notion of 'model' by distinguishing between a *neural generative* model and an *organismic generative* model. Generative models are probabilistic models. They are probabilistic models of the co-occurrence of causes of input and the actual sensory input the brain receives (over time), from which the brain can generate predictions. Now, neurally encoded generative models might very well be the beginning of an account of perceiving but they need not be the ends to that account. To see this, consider the notion of an organismic model present in the following passage from Friston and colleagues:

> We must here understand 'model' in the most inclusive sense, as combining interpretive dispositions, morphology, and neural architecture, and as implying a highly tuned 'fit' between the active, embodied organism and the embedded environment (2012a, p. 6).

Failing to consider this second use of 'model,' highlighting only the neurally encoded generative model is what clears the way for the view that the brain is a virtual reality generator. For it is the notion of the predictive *brain*—secluded from the hidden causes of the environment—that is part of the underlying vision of the brain as a virtual reality generator. But it is plausible to suggest—given the notion of an organismic generative model—that minimization of prediction error is not restricted to the brain alone but involves the entire organisms (morphology, action capacities, and so on) and its embedding environment. It is due to this extended sense of 'model' that Friston claims that an "agent does not *have* a model of its world—it *is* a model." (2013, p. 213; see also Bruineberg and Rietveld 2014; Kirchhoff 2015a, b)

This broader sense of model seems like a good candidate by which to free ourselves from the idea that the brain is a virtual reality generator given that the way we come to perceive the world is not wholly determined by neurally encoded generative models but is in part realized by a much wider set of elements such as body morphology, regularities in the environment, and the particular kinds of sensory organs that we have (see also Clark forthcoming). So by 'model' it does not follow that an organism has an internal, representational model of its niche and that it is this model that does all the cognitive work (if you like). Instead, an organism *is* a model, viz., the causal and statistical regularities reflected in some environment are reflected in some phenotype, i.e., model. In this sense, an organism or phenotype is an optimal (if approximate) model of its environment (Friston 2011, p. 89).

Before we adopt this conclusion, we need first deal with one suspicion that some readers might have: that saying that an organism *is* a model, and from this inferring that it is the entire organism that is involved in minimizing prediction error is equivalent to saying that organisms have long-term priors. If this worry goes through, we are back with having to appeal to priors, and it is not clear that we have moved beyond internalism.

Is this a serious problem? I do not think so, and for a reason to do with how to unpack the notion of 'prediction' (and 'inference'). Following Anderson and Chemero (2013), consider two different senses of 'prediction'. As they say: "The first sense of "prediction" (henceforth prediction$_1$) is closely allied with the notion of correlation, as when we commonly say that the value of one variable "predicts" another (height predicts weight; education predicts income; etc.)." (2013, p. 204) Prediction$_1$ does not involve recruiting hypotheses or, for that matter, testing hypotheses against one another. What Anderson and Chemero term prediction$_2$, however, is "allied … with abductive inference and hypothesis testing." (2013, p. 204) It is this second kind of predictive process that involves "inferring the (hidden) causes of our current observations, and using that hypothesis to predict future observations" (2013, pp. 204–205). Now, a common observation is that to predict something in the environment requires having certain assumptions about the environment. Yet, assumptions can be articulated in terms that do not imply a hypothesis-rich formulation such as when the wavelength selectivity of photoreceptors reflects assumptions about the wavelength of ambient light (see Friston et al. 2012b, p. 5). This suggests a 'minimal' notion of prediction (and also inference) by which it is coherent to say that for any two dynamical systems, A and B, it is coherent to say that A 'models (or 'infers') the "hidden causes of its input (the dynamics of B) when it reliably covaries with the dynamics of B and it is robust to the noise inherent in the coupling." (Bruineberg and Rietveld 2014, p. 7) This is true in the case of prediction$_1$ but not for the hypothesis-rich notion of prediction$_2$. So, and consistent with the view that 'model' implies a highly tuned *fit* between organism and environment, it follows that the statistical and causal structure of the environment comes to reflect the causal and statistical structure of the organism—where 'organism' is understood in the specific sense of being a model of its niche.

There is an intuitive way to elaborate this point further. We can ask: if one appeals to Bayes, where do priors (i.e., the hypotheses) come from (Hohwy 2012)? Approximate empirical Bayes is one way of providing an answer. Here the basic idea is that priors are unearthed from hierarchical statistical learning (Hohwy 2012). According to Hohwy, the "notion of hierarchical inference is crucial here, [since] it enables the brain to optimize its prior beliefs on a moment to moment basis." (2012, p. 3) Some priors are formed in virtue of exposure to causal-statistical patterns in the sensory stimuli over time (e.g. developmental time). Yet others are likely to be more "hard-wired and instantiated over an evolutionary time-scale." (Hohwy 2012, p. 3) Such priors need not be neurally instantiated but rather reflected in our bodily biomechanics, bodily shape, and so on. If correct, it lends support to the possibility that brains do not infer virtual realities. This is so because it is not merely the brain but rather a wider brain–body-world dynamic that realizes prediction error minimization.[5]

Once we acknowledge that the whole organism restricts the uncertainty of the inference space in virtue of an evolved fit between organism and niche, we arrive at the following picture. Embodied and situated activity is part of what enables an organism to restrict the space of probability distributions. This gives us one possible reason to be skeptical about the claim that the brain is a virtual reality generator precisely because what we perceive comes to reflect our embodied model that "distils and embodies causal structure in its local environment." (Friston 2011, p. 89; see also Clark 2016)

So far we have been primarily concerned with perception in discussing the PP scheme. Reflection on the distinction between a neurally encoded generative model and an organismic generative model led to the claim that the internal, generative model is a necessary but not sufficient part of a wider realization base for prediction error minimization. If this is correct, then prediction error minimization is achieved not simply by revising a neurally encoded generative model via perception but depends on a broader set of features such as body morphology and our particular kinds of sensory receptors. This fits naturally with another key part of the PP story: *active inference*, or *action*. The basic idea is that rather than revising its generative models such a system can keep its model constant and instead seek to change the input by selectively sampling the world (Friston 2009; Hohwy 2012).

What is revealing about the addition of active inference to PP is that *without* active inference a system would not be able to minimize prediction error (Hohwy 2012). Proponents of PP often appeal to a tight coupling between perception and action (Clark 2016; Friston 2011; Hohwy 2013). But closer inspection reveals that perception and action are not only tightly coupled but that action is a precondition for perception. In active inference, perception is embodied because perceiving is an activity in which perception and action are constitutive intertwined (for a similar view, see O'Regan and Noë 2001; Noë 2004, 2009);[6] and it is counterfactual in the sense that perception involves patterns of regularities relating sensory input to movement and change (Seth 2014). Crucially, here action and perception are not merely depicted as coupled to one another. Instead, action is an inherent part of perception. In this sense, active inference is consistent with the enactive vision of perception according to which perceiving is widely realized in a brain–body-world dynamics (Noë 2004). All this suggests that active inference puts the brain in a position to utilize embodied activity. Since embodied activity is a strategy for learning about and exploring long-term regularities in the environment, active

[5] Further support for this view comes from the idea that organisms embed the regularities of their embedding niche into their anatomy. Friston and Stephan (2007) puts this in the following way: "the environment unfolds in a thermodynamically structured and lawful way and biological systems embed these laws into their anatomy" (2007, p. 422).

[6] The claim that perception and action are constitutively intertwined is controversial in light of the causal-constitutive distinction, and the alleged causal-constitutive fallacy, in the recent discussions over the extent of mind. However important, I shall not discuss this fallacy here given that it would take us to far astray (see Kirchhoff (2015a, c) for a discussion of the causal-constitutive fallacy in the context of the extended mind debate and the free energy principle).

inference is part of the reason for why the brain is able to settle on a hypothesis with the highest posterior probability. It is a solution that makes PP models continuous with insights concerning world-involving cognition stressed by embodied, enactive accounts of mind. For enactivists, e.g. active engagement in the world "effect[s] statistical regularities that shape the structure of the nervous system" (Gallagher et al. 2013, p. 422).[7]

At this stage in the argument it is worth considering a final worry, for by doing so we can add a final aspect to the claim that action is an inherent aspect of perception. The worry goes as follows: given that the shopkeeper forms an inaccurate prediction about 'fork handles' this suggest that what the shopkeeper perceives is akin to a virtual reality after all; and if so, does this not (contra the EE view) establish that perceiving is at one remove from the world? But this worry is no threat to the EE view. To see this, the central feature to consider is that action does not take place in isolation but is always embedded in specific cultural practices (Roesptorff et al. 2010). According to Roepstorff et al. from "the inside of a practice, certain models of expectancy come to be established, and the patterns, which over time emerge from these practices, guide perception as well as action." (2010, p. 1056) If we apply this idea to the case of the shopkeeper, what we arrive at is not best understood in terms of virtual realities—at one remove from the world—but of perceiving as directly shaped by regularities exhibited in the practices of the shopkeeper's daily life. In other words, in order to explain that she perceives, we cannot 'throw away the world'—as suggested by the inferred fantasies view—but must make direct reference to the particular patterns of practices in which she is embedded and continuously engaged.

Thus far we have considered several aspects of the EE-view, which sets it apart from the inferred fantasies view. First, we have noted that *priors* within the PP scheme need not be neurally instantiated but may be realized extra-neurally in the morphology and action tendencies of the organism. Second, active inference—when understood through the lens of embodied, enactive cognition—suggests that action is in perception. Finally, action does not take place in isolation but unfolds in cultural practices, which shape perception and action. If these aspects turn out to be correct, they indicate (in contrast to the inferred fantasies view) that perception is not at one remove from the environment but is realized in dynamics distributed across the organism and its embedding environment.

5 Conclusion

What does all of this tell us about the relationship between perceiving and imagining? Here I sketch a plausible answer. Both the inferred fantasies view and the EE view states that it is in virtue of prediction error minimization that perceivers are imaginers too. As Clark (2016) puts it: "they are creatures poised to explore and experience their world not just by perception and gross physical action but also by

[7] See also Chemero (2009), Clark (2008), Gibson (1979), Menary (2007), and Thompson (2007).

means of imagery, dreams, and (in some cases) deliberate mental simulations." (2016, p. 84) There are at least two ways in which one can understand the similarities between perceiving and imagining, according to both the inferred fantasies view and the EE view. The first is that both phenomena come about as a result of the (roughly) same underlying process of ongoing prediction error minimization. The second is that in both perceiving and imagining it is the prior probabilities that comprise the generative model that form the basis of top-down predictions.

These similarities between perceiving and imagining also point to ways in which they come apart. Consider what Clark (2016) says about the role of generative models in PP: "the knowledge (model) encoded at an upper layer [in the processing hierarchy] must be such as to render activity in that layer capable of predicting the response profile at the layer below … Since this story applies all the way down … that means that such systems are fully capable of generating 'virtual' versions of the sensory data for themselves." (2016, p. 93) The claim that I defend here is that 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. This is to concede the point that imagining is a neurally realized process in the proximate here-and-now. However, this is not a challenge to the EE view given that it can accept that imagining is an internally realized phenomenon and still hold on to the claim that perception is embodied and world-engaging. Otherwise put, imagining need not involve the brain trying to cancel out error signals elicited (in part) by bodily feedback and environmental input but "rather a kind of multilevel virtual analogue of the driving sensory signal" (Clark 2016, p. 85). This is consistent with the general view of imagining as simulations or recreations of real world perception (Currie and Ravenscroft 2002). Thus, whereas perception is directly constrained and shaped by bodily and environmental elements, this does not hold for imagining. It is this difference that motivates the claim that even if both phenomena are the result of minimizing prediction error, perceiving may nevertheless be widely realized. This is consistent with something akin to the received view of the relation between perceiving and imagining. Thus, even if perceiving and imagining are both realized in prediction error minimization, this does nothing in itself to threaten the received view that imagining and perceiving are different in kind.

Acknowledgements This work was supported by an Australian Research Council Discovery Project "Minds in Skilled Performance" (DP170102987), a John Templeton Foundation Grant "Probabilitizing Consciousness: Implications and New Directions", by a Mind, Brain and Cognitive Evolution fellowship at Ruhr University Bochum, and by a John Templeton Foundation Academic Cross-Training Fellowship (ID #60708). The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the John Templeton Foundation. Thanks to Julian Kiverstein, Jelle Bruineberg, Erik Rietveld, Micah Allen and Jon Opie for comments on a previous version of this paper. Thanks also to the audience members of the Acting Ahead of Actuality conference at the University of Dubrovnik, Croatia, 17-18 June 2016, and to the audience members of the Imagination and Representation workshop at Flinders University, Adelaide, 28 September 2015, for valuable suggestions for improvement. And finally thanks to an anonymous reviewer.

References

Anderson, M., & Chemero, A. (2013). The problem with brain GUTs: Conflation of different senses of "prediction" threatens metaphysical disaster. *Behavioral and Brain Sciences, 36*(3), 204–205.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11*(7), 280–289.

Bruineberg, J., & Rietveld, E. (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances. *Frontiers in Human Neurosciences*. doi:10.3389/fnhum.2014.00599.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: The MIT Press.

Clark, A. (2008). *Supersizing the mind: Action, embodiment, and cognitive extension*. New York: Oxford University Press.

Clark, A. (2012). Dreaming the whole cat: Generative models, predictive processing, and the enactivist conception of perceptual experience. *Mind, 121*(483), 753–771.

Clark, A. (2013a). Expecting the world: Perception, prediction, and the origins of knowledge. *The Journal of Philosophy, CX*(9), 469–496.

Clark, A. (2013b). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36,* 181–253.

Clark, A. (2014). Perceiving as predicting. In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its modalities*. New York: Oxford University Press.

Clark, A. (2015). Radical predictive processing. *The Southern Journal of Philosophy, 53,* 3–27.

Clark, A. (2016). *Surfing uncertainty*. Oxford: Oxford University Press.

Clark, A. (forthcoming). *Busting out*: *Predictive brains, embodied minds, and the puzzle of the evidentiary veil*. Downloaded from: http://www.researchgate.net/publication/287216594.

Colombo, M., & Wright, C. (2016). Explanatory pluralism: An unrewarding prediction error for free energy theorists. *Brain and Cognition*. doi:10.1016/j.bandc.2016.02.003.

Currie, G., & Ravenscroft, I. (2002). *Recreative minds: Imagination in philosophy and psychology*. New York: Oxford University Press.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 360,* 815–836.

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Cell Press*. doi:10.1016/j.tics.2009.04.005.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127–138.

Friston, K. (2011). Embodied inference: Or 'I Think Therefore I Am, If I Am What I Think'. In W. Tschacher & C. Bergomi (Eds.), *The implications of embodiment (cognition and communication)* (pp. 89–125). Exeter: Imprint Academic.

Friston, K. (2013). Life as we know it. *Journal of the Royal Society, Interface, 10,* 20130475. doi:10.1098/rsif.2013.0475.

Friston, K., Adams, R., & Montague, R. (2012a). What is value-accumulated reward or evidence? *Frontiers in Neurorobotics*. doi:10.3389/fnbot.2012.00011.

Friston, K., & Stephan, K. E. (2007). Free energy and the brain. *Synthese, 159,* 417–458.

Friston, K., Thornton, C., & Clark, A. (2012b). Free-energy minimization and the dark-room problem. *Frontiers in Psychology, 3*(130), 1–7.

Gallagher, S., Hutto, D. D., Slaby, J., & Cole, J. (2013). The brain as part of an enactive system. *Behavioral and Brain Sciences, 36*(4), 421–422.

Gendler, T. (2011). Imagination. *Stanford Encyclopedia of Philosophy*, 1–32.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. New York: Taylor & Francis.

Hobson, J., & Friston, K. J. (2014). Consciousness, dreams, and inference: The Cartesian theatre revisited. *Journal of Consciousness Studies, 21*(1–2), 6–32.

Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Frontiers in Psychology*. doi:10.3389/fpsyg.2012.00096.

Hohwy, J. (2013). *The predictive mind*. Oxford: Oxford University Press.

Hohwy, J. (2014). The self-evidencing brain. *Nous*. doi:10.1111/nous.12062.

Hohwy, J. (2015). The neural organ explains the mind. In T. Metzinger & J. Windt (Eds.), *Open MIND, 19(T)*. Frankfurt am Main: MIND Group. doi:10.15502/9783958570016.

Hohwy, J., Roepstorff, A., & Friston, K. (2008). Predictive coding explains binocular rivalry: An epistemological review. *Cognition, 108*(33), 687–701.

Kirchhoff, M. D. (2015a). Extended cognition & the causal-constitutive fallacy: In search for a diachronic and dynamical conception of constitution. *Philosophy and Phenomenological Research, 90*(2), 320–360.

Kirchhoff, M. D. (2015b). Experiential fantasies, prediction and enactive minds. *Journal of Consciousness Studies, 22*(3–4), 68–92.

Kirchhoff, M. D. (2015c). Species of realization and the free energy principle. *Australasian Journal of Philosophy, 93*(4), 706–723.

Kosslyn, S. M. (1994). *Image and mind*. Cambridge, MA: Harvard University Press.

Menary, R. (2007). *Cognitive integration: Mind and cognition unbounded*. Basingstoke, UK: Palgrave Macmillan.

Noë, A. (2004). *Action in perception*. Cambridge, MA: The MIT Press.

Noë, A. (2009). *Out of out heads*. New York: Hill & Wang.

O'Regan, K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939–973.

Orlandi, N. (2014). *The innocent eye: Why visio is not a cognitive process*. Oxford, NY: Oxford University Press.

Paton, B., Skewes, J., Frith, C., & Hohwy, J. (2013). Skull-bound perception and precision optimization through culture. *Behavioral and Brain Sciences, 36*(4), 222.

Rescorla, M. (2013). Bayesian perceptual psychology. In M. Matthen (Ed.), *The Oxford handbook of the philosophy of perception*. New York: Oxford University Press.

Roesptorff, A., Niewöhner, J., & Beck, S. (2010). Enculturating brains through patterned practices. *Neural Networks, 23,* 1051–1059.

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.

Stevenson, L. (2003). Twelve concepts of imagination. *British Journal of Aesthetics, 43*(3), 238–259.

Stone, J. V. (2013). *Bayes' rule*. Sheffield: Sebtel Press.

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Cambridge, MA: Harvard University Press.

Ward, D. (2012). Enjoying the spread: Conscious externalism reconsidered. *Mind, 121*(483), 731–752.