

Penultimate draft, to appear in *Mind & Language*

## **The rational role of the perceptual sense of reality**

**Paweł Gładziejewski**

**Department of Cognitive Science**

**Nicolaus Copernicus University in Toruń, Poland**

### **Correspondence**

Paweł Gładziejewski

Department of Cognitive Science, Nicolaus Copernicus University in Toruń, Gagarina 39, 87-100, Toruń, Poland

Email: pawel.gla@umk.pl

### **Funding information**

National Science Center in Poland, grant number 2019/33/B/HS1/00677

Perceptual experience usually comes with “phenomenal force”, a strong sense that it reflects reality as it is. Some philosophers have argued that it is in virtue of possessing phenomenal force that perceptual experiences are able to non-inferentially justify beliefs. In this paper, I introduce an alternative, inferentialist take on the epistemic role of phenomenal force. Drawing on Bayesian modeling in cognitive science, I argue that the sense of reality that accompanies conscious vision can be viewed as epistemically appraisable in light of its rational etiology.

### **KEYWORDS**

dogmatism, perceptual justification, Bayesian brain, predictive processing, reality monitoring, derealization

### **1. Introduction**

Part of what it is like to consciously see a visual scene—for most human perceivers, under usual circumstances—is to experience the objects that comprise it as though *they are really there right now*. Visual experience involves the feeling of being immersed in the world or being

accurately presented with one's environment. I will refer to this aspect of visual phenomenology as "phenomenal force" (henceforth PF) or "perceptual sense of reality".

It is notoriously hard to pin down what PF consists in, exactly. In fact, one of the aims of the present paper is to attempt to elucidate this notion. For now, it might be heuristically helpful to list some contrast cases, that is, perceptual or perceptual-like states that arguably *lack* PF. Take these three types of examples.

First, consider the state of seeing "stars" after hitting yourself in the head, or seeing eye floaters when looking at a bright sky with your eyes squinted, or the state undergone by a synesthete who, upon hearing b-minor, begins to see as a two-dimensional color patch overlaid on a region of her visual field. Each of those states possesses visual sensory phenomenology. But the trauma-induced stars, eye floaters, or synesthetic color patches are not experienced as present *out there* in the space around oneself, unlike, say, a visually perceived cat or a bouncing basketball.

Second, consider a very vivid visual imagery of a bouncing basketball, and compare it with a visual experience of a basketball bouncing in heavy fog. Introspectively, these states seem to be similar in that they share visual(-like) phenomenology, which may include them being matched in terms of the degree of vividness. Despite the significant phenomenological overlap, perceiving and imagining feel different. Imagining an object, in contrast to perceiving it, is not an experience that feels as though it reveals one's immediate environment as *actually* being a certain way.

Third, take the case of derealization. People who experience derealization report that their perception feels detached, dreamy, comparable to watching a movie. Their visual world feels unreal. Notably, derealization is not associated with an impairment of visual discriminative, recognitional, or action-guiding abilities. Again, what is selectively lacking, in this case, is the sense of the reality of the world as perceived.

Together, these examples establish that it is possible to have a visual or visual-like experience that differs from ordinary visual perception exclusively by lacking one particular aspect of the latter's phenomenology—the sense of reality. Let us assume, then, that this ingredient of normal visual experience constitutes its PF.

My main focus in the present paper is on the epistemic import of PF. Does the PF contribute to the rational powers of visual perception, that is, to its ability to justify beliefs? Assume that I have a visual perception as of a cat and that this visual state justifies my belief that there is at least one mammal in my vicinity. Does the PF of this visual experience, the sense

that it reveals the world as it is, generate or contribute to the justification for my belief? If so, then how?

Here is a seemingly plausible take on this. Over the past two decades or so, some philosophers have opted for a “dogmatist” or “conservative” view of perceptual justification (Chudnoff, 2013; Huemer, 2001; Pryor, 2000; Tucker, 2010). At its core, this position combines two claims. One claim is that perceptual justification works the way it pretheoretically seems to work (Pryor, 2000). Simply, when a subject undergoes a visual experience with the content *p*, then, absent defeaters, the subject is justified in believing that *p*. The second claim is that perception provides immediate or non-inferential justification. That is, perceptual experiences do not owe their justification-conferring powers to any inferential connections to preexisting mental states.

Now, any sort of foundationalist view like this faces the following problem (BonJour, 1985; Sellars, 1956). Dogmatists take perceptual states to have intentional content. It seems, however, that the ability of a contentful mental state to confer justification requires this state to be itself (inferentially) justified by other, antecedent contentful states. How can perceptual states be excused from having to obey this rule? How can they transfer justification to beliefs without needing to be themselves justified? This is where the notion of PF comes in. Perhaps the special status of perceptual states is somehow grounded in their distinct, world-revealing phenomenology. Intuitively, when you consciously see a cat, it makes no sense to ask whether you are justified in having your visual state. The cat just feels immediately present *right there*.

Plausible as it may seem, I think we should be apprehensive about this sort of position. One concern is that instead of solving the problem of unjustified justifiers, it postpones it. A legitimate question can still be raised as to what it is about perceptual states *that possess PF* that allows them to confer justification without needing to be justified. From a purely psychological standpoint, it may well be that when a perceptual state possesses PF, the perceiver does not experience the need for having further justification, nor shows behavioral tendencies to seek such justification. But this descriptive point is irrelevant here. What we are after is some principled *normative* story that renders the perceiver rationally entitled to her belief in virtue of her perceptions having PF.<sup>1</sup>

---

<sup>1</sup> Perhaps we should simply claim that PF only plays such non-normative role. Later in this paper, I will argue that, as matter of fact, the PF has a rational etiology. We thus have a reason for treating PF as epistemically appraisable that is independent from considerations regarding foundationalism.

Another observation to make here is that the PF of perceptual experience is generated by fallible cognitive mechanisms (Ghijssen, 2014; Teng, 2016; Siegel & Silins, 2015). These mechanisms operate beyond conscious awareness, and so their existence may not be readily apparent for most neurotypical subjects during everyday life. Still, empirical research suggests that these mechanisms can be erroneous in epistemologically interesting ways. The false-positive errors happen when a mental state ill-suited to play an epistemic role of perception is experienced as a (PF-possessing) perception. This category includes, for example, perceptual reality monitoring errors, where endogenous imagery is misidentified as a perceptual state, resulting in a hallucination (Johnson, Hashtroudi, & Lindsay, 1993; Simons, Garrison, & Johnson, 2017; I will return to this idea in Section 2.2). The false-negative errors include cases where a state that is well-suited to play an epistemic role of perception (because it *is* a perception) fails to be experienced as world-revealing. This is illustrated (at least on one controversial interpretation, see Teng, 2016) by a finding, due to Perky, that when subjects are asked to visually “project” an imagined banana on a screen while a faint picture of a banana is actually presented unbeknownst to them, they seemingly conflate their perception for imagery (Perky, 1910; for more recent research that substantiates the idea of false-negative errors of this kind, see Dijkstra & Fleming, 2021; Dijkstra et al., 2021).

Perhaps the nature of the processes at play behind the production of the PF of perceptual experience could shed light on the involvement of PF in perceptual justification? Perhaps what contributes to the rational role of PF is its etiology? What I mean to suggest is that PF *itself* may be epistemically evaluable based on whether its etiology is epistemically appropriate, and its ability to confer justification (or contribute to the justification conferral) on beliefs may depend on whether it has been generated in an epistemically appropriate manner.

The main aim of this paper is to establish such a view as a viable theoretical option. I will draw on Bayesian modeling in cognitive science to show how PF could be generated by unconscious probabilistic inference. If this etiological story is right, then it is being grounded in inference that allows the PF to play a rational role in cognition. Thus, the architecture of PF-based perceptual justification is not foundationalist, but inferential.

The discussion is structured as follows. I start, in Section 2, by investigating the question of how we should construe the PF of visual experience. I argue that there are three distinct and independent aspects of visual experience that contribute to the overall sense of reality. The three respective aspects are related to: (1) the objectification of first-order perceptual content; (2) perceptual reality monitoring of the first-order content; (3) forming interoceptive appraisals of visual contents.

In Section 3, I draw on Bayesian view of perception—in particular, on predictive processing—to show how each of the three aspects of visual perception mentioned above can be regarded as corresponding to a hypothesis generated through unconscious Bayesian inference. This establishes a rational relation between the PF, construed as a conclusion of inference, and internally stored prior assumptions. A picture thus emerges according to which the ability of the PF-possessing perceptual states to confer justification is grounded in the rational etiology of the PF itself.

In Section 4, I take a step back to address some general concerns about my proposal. I end the paper with brief concluding remarks.

## **2. Phenomenal force, three ways**

In this section, I introduce the idea that there are three distinct and independent aspects of visual experience, each of which contributes to the sense of reality. To elucidate those three aspects, I investigate their contrast cases, that is, examples of mental states that lack one of the aspects but are much like normal perception otherwise.

### **2.1. Phenomenal force as perceptual objectification**

Let us start by revisiting the first batch of cases mentioned at the outset: seeing “stars”, eye-floaters, and synesthetic color patches. Other examples of this kind may involve seeing one’s whole visual field filled with a uniform color or seeing objects depicted on a two-dimensional picture, like a painting (Matthen, 2010). It seems that all those cases lack a certain feature shared by the states of seeing scenes comprised of ordinary physical objects: a basketball bouncing on the floor, a vase full of tulips, a pair of dogs playing in the garden. The notion of perceptual objectification aims to capture the difference-maker at play here.

Each of the mental states mentioned involves the visual experience of features like color, size, shape, or texture. On a plausible first approximation, what distinguishes seeing ordinary objects from the contrast cases lies in the way the visual features are organized in the experience. In ordinary perception, the visual features are organized in such a way that the content of a perceptual state exhibits *objectivity*, in that it purports to represent “the physical environment, beyond the individual’s local, idiosyncratic or subjective features” (Burge, 2010, p. 397). Ordinary objects seem to be located out there, in a way that transcends our perspectival

takes on them. The contrast cases lack this perspective-independent presence altogether or possess it to a significantly diminished degree.

Another, closely related take on the relevant difference is this: In ordinary perception, the visual features are bounded together in a way that yields conscious visual experiences as of objects that *have* those features. For example, a series of visual features may be experienced as belonging to an object. Such object is perspective-independent by being experienced as retaining its identity and properties against the shifts in the perceiver's perspective. For example, a basketball may be perceived as the same single thing that retains its size even as it appears larger in the visual field when the perceiver approaches it.<sup>2</sup>

To get a still clearer view of perceptual objectification, let us follow the idea that the capacity for ordinary object-seeing is closely related to the *predictive* structure of vision (Church, 2013; Clark, 2012; Madary, 2017; Noë, 2002; Seth, 2014; Siegel, 2010). Underlying this view is a notion that the content of a perceptual state is determined, at least in part, by expectations regarding how the experience will change as a result of the shifting of one's perspective. The idea, then, is that a perceptual state possesses objectified content when its content is determined (or co-determined) by a specific pattern of sensory predictions. In particular, four types of predictions seem crucial:

- (1) Property constancy-in-variation predictions—seeing an ordinary object involves attributing to it factual properties, which in turn consists in predicting a pattern of perspective-dependent changes in the way the properties appear. For example, perceiving an object as possessing a factual property of being round may involve the prediction that it will gradually turn from appearing round when being looked at straight, to appearing elliptical when looked at under an angle.
- (2) Perspectival completion predictions—seeing an ordinary object involves the prediction that as one moves with respect to the object, previously unseen sides of the object will be revealed. It is in virtue of having such predictions that we can be said to visually experience a whole, three-dimensional object, rather than just its front part.

---

<sup>2</sup> This point is meant to capture generic phenomenology of seeing objects. As such, it remains uncommitted with respect to a more technical question of whether the objects tracked in visual perception are best modeled as bundles of properties, or as bare particulars to which the properties are attributed, or in terms of some other ontological category altogether (for some relevant work, see Skrzypulec, 2017, Vernazzani, 2021).

- (3) Self-movement-independence predictions—seeing an ordinary object involves the prediction that its properties are independent from the perceiver’s movement (see Siegel, 2010, Ch. 7). For example, this prediction is violated if an object remains “glued” to the same position in your visual field as you change perspective (through the eye- or head movement).
- (4) Object-interaction predictions—seeing an ordinary object involves predicting that the visual experience will change as a function of the causal and spatial relations between the object in question and other objects that comprise the visual scene. For example, visually experiencing a fence as being located in front of a tree involves the expectation that when one moves relative to the scene, parts of the fence will occlude the tree and not the other way around.

The claim, then, is that perceptual objectification is the ingredient missing in all the contrast cases. This point can in turn be cashed out in terms of the contrast cases lacking the predictive structure of normal visual experience. Take the most obvious case of the experience of a uniform color filling the whole visual field. Here, there is no object to which the color is attributed, and the experience fails to fulfill any sensory predictions in response to the changing perspective of the subject.

The cases of seeing stars, seeing a sound-triggered patch of color, or seeing eye-floaters are closely similar. Although a synesthetic patch of red may exhibit minimal objecthood simply by being spatially bounded and cohesive, it falls short of meeting the predictive structure of ordinary object perception. For example, the color of the patch does not exhibit constancy-in-variation as the ambient illumination changes; trying to move around the patch fails to reveal its backside; it remains locked in the same position in the visual field as the perceiver shifts her gaze; it behaves “ghostly” by failing to interact with physical objects.

The remaining case of picture-seeing is slightly different from the other examples. Seeing objects depicted on a flat surface retains some of the predictive structure of ordinary object perception. In particular, depicted objects exhibit some degree of constancy-in-variation (for example, with respect to color) and they remain independent of the perceiver’s movement. Still, there are significant differences to full-blown ordinary object perception, as depicted objects fall short of meeting perspectival completion predictions and the predictions regarding the spatial and causal interactions with other objects. Although picture-seeing is closer to normal object perception than the other contrast cases, it still differs significantly.

## 2.2. Phenomenal force as perceptual reality monitoring

Objectification of perceptual content cannot be the whole story about the PF of visual perception. Note that humans can run stimulus-free simulations of visual experience. At least some such simulations are conscious and their content exhibits objectification. A striking example is false awakenings that occasionally follow lucid dreams. Immediately after experiencing a lucid dream, a person may undergo an extremely realistic *non*-lucid dream in which she apparently wakes up in her room and gets on with her day (Windt, 2015). It is reasonable to assume that false awakenings involve perceptual objectification: the subject visually experiences being surrounded by ordinary objects in a way that closely corresponds to waking experience.

Of course, false awakenings (and non-lucid dreams in general) feel veridical, and thus do not constitute a contrast case for states that possess PF. But consider a more prosaic example of visual imagery. Just as we may dream of walking through a room, we may vividly imagine doing so. Imagining seeing a cat lying on a floor may mimic the predictive structure of ordinary vision (say, when imaginative investigation reveals a previously occluded side of the imagined cat). Yet, there is a crucial phenomenological difference between seeing and visualizing a cat. Perceiving a cat feels as though one is confronted with an actual cat. Imagining a cat does not feel like being confronted with anything in one's actual surroundings. Because the contents of both states are objectified, perceptual objectification cannot account for the difference between them. The difference I am after here also cannot be fully cashed out in terms of the differing degrees of vividness or level of detail. Imagery can sometimes match visual perception in those respects (for example, in people who experience hyperphantasia; see Zeman et al., 2020) while still failing to feel real in the relevant sense.

My proposal is that the crucial phenomenal ingredient that imagery lacks is not to be found at the level of first-order perceptual content. Rather, it has to do with a metacognitive assessment of or attitude towards the first-order content (see also Dokic & Martin, 2017). Normal visual perception is usually accompanied by a metacognitive sense (or feeling) of the first-order content as confronting the subject with a worldly state of affairs. On an alternative formulation, perception can be thought of as being associated with a metacognitive attitude towards the first-order content as accurately representing the current state of the environment. Visual imagery is usually associated with a metacognitive sense (or feeling) of its first-order



content as being endogenous in origin, or a metacognitive attitude towards the first-order content as representing a past, future, or counterfactual (rather than actual) state of affairs.<sup>3</sup>

I think that the metacognitive construal fits some of the existing attempts to capture the nature of PF. For example, Huemer writes:

Forcefulness is not a matter of either the qualia or the representational content of an experience; it is a third aspect of experience. ... it is the fact that, in the experience, it seems to one that something satisfying the content of the experience actually exists, here and now (Huemer, 2001, p. 79).

This third aspect to which Huemer refers is, on the present view, metacognitive. Relatedly, it seems that something like the metacognitive construal is suggested by Pryor when he claims that experience represents certain propositions in such a way that these propositions “feel as if” we can tell they are true just by having them so represented (Pryor, 1999, p. 547), or by Tucker when he describes the phenomenology of a (PF-possessing) mental state in terms of the state “recommending” its propositional content as true (Tucker, 2004, p. 530). Those ways of describing PF are, I think, naturally interpreted as pointing to a metacognitive evaluation of the first-order content.

The metacognitive view gains traction when we focus on theoretical and empirical considerations from cognitive science. There are functional constraints that give leverage to the idea that metacognition should play a role in perception. Perception and imagery are subserved by a largely shared neural machinery (for a review, see Pearson, 2019). This arguably creates a need for a “reality monitoring” mechanism that would allow the brain to distinguish internally-generated activations of visual areas from activations caused by the external environment—a mechanism for avoiding false-awakening-like scenarios, so to speak (Dijkstra, Kok, & Fleming, 2021; Gershman, 2019; Lau, 2019). More specifically, the brain would need to keep apart three types of cases: (1) random activations of perceptual areas (noise) that do not serve a cognitive function; (2) endogenous activations of perceptual areas that subserve non-perceptual functions, like imagery, prospection or episodic memory; (3) exogenously-driven activations of perceptual

---

<sup>3</sup> It needs stressing that the metacognitive sense, feeling or attitude that I am trying to capture here is *not* equivalent to a belief. Staunch solipsists still enjoy perceptual experiences that subjectively *feel* real or veridical. On the present proposal, the metacognitive aspect of PF is integral to the phenomenology of perceptual experience itself.

areas that subserve perception (Lau, 2019). If this is right, then for a state to possess the metacognitive aspect of PF is for it to be estimated by the reality monitoring mechanism as belonging to the last category.

There is substantial, even if indirect empirical evidence that supports the existence of such perceptual reality monitoring mechanism (the discussion that follows is based on: Simons, Garrison, & Johnson, 2017). Perceptual reality monitoring is thought to be an instantiation of a more general process of reality monitoring, which additionally includes the ability to distinguish memories (exogenously caused by a past event) from imagined events. Of particular interest here are studies on people who suffer from schizophrenia, many of whom experience hallucinations. It has been repeatedly found that people who suffer from schizophrenia tend to produce more memory-related reality monitoring errors than healthy individuals when distinguishing previously seen from previously self-generated stimuli. Furthermore, this pattern of errors exhibits an externalization bias, whereby a disproportionately large amount of errors consist of internally generated contents misidentified as having been externally generated. The externalization bias for memories is especially pronounced in those patients who experience hallucinations. Furthermore, the tendency to experience hallucinations among people with schizophrenia is associated with the disruption of the activity of the anterior prefrontal cortex, an area that has been independently found in healthy individuals to be involved in reality monitoring tasks for memory.

On a plausible interpretation of those results, hallucinations involve a sort of reality monitoring error where the endogenous neural activity in perceptual areas is misidentified as exogenous, and thus obtains the sense of reality that normally characterizes perception. Errors of this sort are perceptual equivalents of the externally biased reality monitoring errors for memory and stem from the same core metacognitive mechanism.<sup>4</sup>

One last point regarding the relationship between the reality monitoring and the perceptual objectification aspects of PF. As already noted, a state that lacks the metacognitive aspect can still possess the objectification aspect. But reverse cases—states that exhibit the metacognitive PF but lack, at least to some relevant degree, objectification—are also possible. Consider again seeing objects depicted on a painting. Seeing a cat depicted on a flat surface does not feel like imagery. Rather, the metacognitive sense (or feeling, or attitude) that accompanies this state is that of normal visual perception: the state feels like being perceptually

---

<sup>4</sup> The Perky finding mentioned in the introduction could, from this perspective, constitute an internally biased error, whereby a perceptual state is erroneously estimated as endogenous in origin.

confronted with something actual. However, because the part or aspect of the state's first-order content that is directed at the depicted cat lacks objectification<sup>5</sup>, it fails to deliver an object of the sort that one could be "confronted with" in a fully immersive manner. Thus, the cat-directed aspect of the state feels perceptual, but is not objectified. We may thus conclude that perceptual objectivity and reality monitoring are separate and independent aspects of PF.

### 2.3. Phenomenal force as interoceptive appraisal

Derealization constitutes the remaining case of visual experience devoid of PF. It consists of feeling as though one is cut off from one's perceptual world, or the feeling that the world is unreal or foreign. In trying to flesh out the nature of derealization, we may start by noticing its close connection to depersonalization. The two conditions are so tightly associated that they are sometimes treated as two faces of the same underlying disturbance (Medford, 2012; Ratcliffe, 2008). If derealization consists of feeling cut off from the world, depersonalization involves the feeling of being cut off from oneself, as expressed by reports of feeling empty, robotic, alienated from one's experience, like an outside observer of one's own mental life.

Let us focus on depersonalization for now. According to a dominant view which I will adopt here, the experience of self-estrangement is explained by de-affectualization (cf. Colombetti & Ratcliffe, 2012; Gerrans, 2019; Medford, 2012). On this approach, depersonalized subjects suffer from significantly reduced emotional reactivity, where the latter is cashed out, along broadly Jamesian lines, in terms of disrupted interoception. Diminished affect is commonly reported by people who experience depersonalization (Medford, 2012). Furthermore, depersonalization is related to the hypoactivity of the anterior insula, an area involved in the processing of bodily signals and in encoding affective salience (Gerrans, 2019; Medford, 2012).

Here is a phenomenological take on the interoceptive view of depersonalization (Colombetti & Ratcliffe, 2012, Ratcliffe, 2008; see also Damasio, 1994, Sass, Parnas, 2003). The crucial idea is that ordinary perceptual experience is accompanied by a continuous interoceptive background state. Although this state consists of subtle bodily feelings, the body is not its intentional object. Instead, the body is experienced as that *through* which the subject

---

<sup>5</sup> Other parts of the state's content (for example, the experience of the picture itself rather than of what is depicted) *do* exhibit objectification. I thank an anonymous reviewer for pointing this out.

experiences the world. The state in question is more mood-like than emotion-like in two respects. First, rather than being bound to a particular context (say, an encounter with a fear-inducing stimulus), it constitutes a continuous and ongoing part of conscious experience. Second, unlike an emotion, it is not intentionally directed at a particular object or trigger (say, a scary clown), but rather “colors” one’s overall experience in an all-encompassing manner. When this interoceptive canvass of experience is disrupted, the lived body (through) which we experience turns into just another worldly object. This is depersonalization.

In this picture, derealization constitutes the world-directed side of disrupted bodily experience. A natural way to understand this is by analogy to the Jamesian view that emotions are bodily feelings through which we appraise the hedonic or motivational value of things we perceive (Damasio, 1994; Prinz, 2006). Again, the sort of interoceptive appraisal that is disturbed in derealization is more long-standing and all-encompassing than the emotional appraisal. It has to do with an ongoing sense of being affectively “at home” in the world, or as being affectively connected to it. It underlies the general sense of the world as practically significant, and the experience of the objects as motivationally salient. When this default affective feeling is disrupted, like in derealization, “the world ceases to move and affect one through one’s body” (Colombetti & Ratcliffe, 2012, p. 148).

This affective connection to reality is distinct from perceptual objectification. People who experience derealization retain the capacity for binding sensory features into perceptual representations of objects and for representing the environment as comprised of such objects.

The interoceptive aspect of PF is also distinct from perceptual metacognition. It would be phenomenologically inaccurate to say that derealized subjects fail to experience perceptual contents as “asserting” (to use a rough analogy to linguistic assertion) that the world is a certain way, or that they experience perceptual contents as endogenous in origin. Although the world feels unreal in derealization, it is still experienced as being *perceptually* encountered rather than *imaginatively* conjured up. Also, if derealization involved disrupted perceptual reality monitoring, it would be natural to expect both conditions to overlap at the level of neural mechanisms, as well as to expect derealization to involve hallucinations akin to those experienced by schizophrenic patients. But neither of those is the case (cf. Sass et al., 2013). More generally, on the present view, the interoceptive side of PF concerns the hedonic or motivational relevance of things to the subject. This evaluative phenomenology comprises *first-order* content, as the relevant evaluations are intentionally directed at the constituents of the world, rather than on the mental representations thereof. The interoceptive side of PF is not metacognitive on the present proposal.

## 2.4. Phenomenal force is complex

One could feel compelled to ask: which of the three features of conscious perception discussed above constitutes its PF? However, I think that the search for “one true PF” is misguided. When the mechanisms of conscious perception operate normally, the three features work in tandem, each contributing to the phenomenology of seeing. It is not easy to disentangle those contributions based on introspection, which presumably inspires an assumption that there is a single aspect of experience that constitutes its PF. But if the considerations above are on a right track, the perceptual sense of reality involves three separate and independent constituents.<sup>6</sup>

Let me unpack this idea using a phenomenal contrast argument. Suppose you perceive a lonely tree on a grassland. Now imagine your experience modified in one of the following ways:

- (1) As you move with respect to the (apparent) tree, your experience fails to change in expected ways. It turns out that what you took for a tree is merely a large, realistic drawing that blends in with the rest of the scene. Still, your state feels perceptual (rather than imagery-like) and you enjoy your normal overall affective connection to your visual world.
- (2) As you move with respect to the tree, you get the sense that what you are experiencing is merely an extremely vivid episodic memory. Still, within this mnemonic phenomenology, the (quasi-)visual expectations are fulfilled and you feel normal affective connection to the scene.
- (3) As you move with respect to the tree, you experience an episode of derealization. You feel affectively numb and estranged from the visual scene. Still, your state feels properly perceptual, and your tree-related perceptual expectations are fulfilled.

---

<sup>6</sup> Of course, if the present paper is on the right track, then these constituents are unified at a deeper, explanatory level, as they share a common core of being underpinned by Bayesian computation. I thank an anonymous reviewer for pointing this out.

Now, case (1) lacks perceptual objectification, case (2) lacks the reality monitoring aspect of PF, and case (3) lacks the interoceptive aspect. The full-blown PF of normal waking visual experience arises when all three components are in place.<sup>7</sup>

From the perspective of normative inquiry, this view of PF complicates things as now we have not one, but three separate aspects of experience whose rational involvement in cognition requires investigation. The upside, however, is that we can turn the question of the epistemic import of PF into a more specific question of the epistemic import of perceptual objectification, perceptual metacognition, and the interoceptive appraisal of visual contents.

### **3. Phenomenal force and Bayesian inference**

In this section, I propose an inferentialist take on the epistemic role of PF. I first outline the Bayesian approach to perception and its epistemological consequences. Then I apply this approach to the three aspects of conscious experience discussed in the previous section.

#### **3.1. Bayesian perceptual justification**

The data streams registered by the sensory apparatus of an organism are inherently ambiguous, as there is no one-to-one mapping that connects the worldly causes with their sensory effects. In this paper, I will follow a historical line of theories that use the underdetermination of sensory states by their causes as a basis for treating perception as an inference problem. The job of the perceptual apparatus, on this approach, is to infer the most likely cause of current stimulation. According to the Bayesian guise of the inferentialist view, this causal inference conforms to Bayes rule:

$$(1) \quad p(h|d) = \frac{p(d|h)p(h)}{p(d)}$$

---

<sup>7</sup> Importantly, although distinct, the three components of PF can interact, either at the level of experience or at the level of cognitive mechanisms that subserve them. For example, interoceptive evaluations are usually directed at objects as represented through objectified contents. To take another example, the reality monitoring system could use interoceptive cues to inform the process of estimating of the source of first-order content. I thank an anonymous reviewer for pressing me on this point.

That is, the computational goal of perception is to find a hypothesis about the causes of the sensory signal that most likely true, given the signal (which acts as raw data,  $d$ ), thus maximizing the posterior probability  $p(h|d)$ . Computing the posterior requires combining the prior belief about the probability of the hypothesis being true,  $p(h)$ , and the likelihood,  $p(d|h)$ , which measures how likely the data is, given the hypothesis.

Here, I will narrow my focus to what has become the dominant variant of the Bayesian approach to perception, namely to predictive processing (PP) (for historical landmarks, see Rao, Ballard, 1999; Friston, 2005; for up-to-date and philosophically oriented introduction and review, see Hohwy, 2020; for a more technical introduction, see Bogacz, 2017). The mathematical idea behind PP is that the brain *approximates* exact Bayesian inference by engaging in what is called “variational” Bayesian inference.

The crux of variational inference lies in the notion that instead of directly calculating the posterior probability distribution  $p(h|d)$ , the brain encodes another distribution,  $q(h)$ , which is iteratively optimized to match the posterior. Formally, the idea is to minimize the Kullback-Leibler divergence (KL-divergence, relative entropy) between  $q(h)$  and  $p(h|d)$ , which measures the difference between the respective distributions:

$$(2) \quad D_{KL}(q(h) \parallel p(h|d)) = \sum_h q(h) \log \frac{q(h)}{p(h|d)}$$

Because the formula above assumes already knowing  $p(h|d)$ , the point is *not* that the brain reduces the KL-divergence directly. Instead, the trick behind PP relies on the following derivation of (2):

$$(3) \quad D_{KL}(q(h) \parallel p(h|d)) = F + \log p(d)$$

Here,  $F$  expresses the quantity of variational free energy:

$$(4) \quad F = \sum_h q(h) \log \frac{q(h)}{p(h, d)}$$

Crucially, given (3), the KL-divergence between the true and approximate posterior can be *indirectly* minimized through minimizing the value of  $F$ . For reasons that will become clear

below, the brain, on the PP view, possesses the resources needed to minimize  $F$ . The idea, then, is that the brain minimizes variational free energy to infer the causes of sensory stimulation.

The next step is to connect the mathematical story of variational inference to neural information processing. PP rests on a postulate that feedback synaptic connections in the brain encode a generative statistical model of the environment. The model aims to capture the structure of causal dependencies between the constituents of the environment and the way this structure generates the sensory input. Formally, the model encodes the joint probability of sensory states and their (postulated) worldly causes,  $p(h, d)$ . Notice, now, that the generative model and the approximate posterior  $q(h)$ —both of which feature in equation (4)—are assumed to be internally encoded and accessible for use in information-processing mechanisms of the brain.<sup>8</sup>

According to PP, the generative model is used to predict the flow of the sensory stream. These predictions are propagated top-down and compared against actual sensory input, giving rise to the prediction error signal. The prediction error is sent up the processing hierarchy and the perceptual hypothesis is readjusted, if needed, to reduce the error.

Through minimizing prediction error, the generative model gets inverted to yield a perceptual hypothesis about the hidden cause(s) of the sensory input. In other words, minimizing the prediction error is equivalent to searching for a  $q(h)$  that best approximates the true posterior under the generative model.

This computational procedure, according to PP, is implemented in causal transactions that underlie perception. This means that prior assumptions encoded in the generative model causally mediate the production of perceptual states. Thus, the transition from sensory states and the assumptions encoded in the model (prior and likelihood distributions) to the output perceptual states is a causal transition that complies with a rational, truth-preserving rule (see also Kiefer, 2017).

A rather unorthodox picture of perceptual justification emerges (Gładziejewski, 2017, 2021a; see also Ghijssen, 2021; Munton, 2018; Vance, 2015). Perceptual states are produced based on their (Bayes-)rational, inferential connections to antecedent mental states. This rational etiology renders them epistemically appraisable (see also Siegel, 2017). Furthermore,

---

<sup>8</sup> Notice also that the joint distribution  $p(h, d)$  that defines the generative model is factorizable into the product of the prior  $p(h)$  and the likelihood  $p(d|h)$ . By solely relying on the prior and likelihood distributions to approximate the posterior, variational inference bypasses the cumbersome process of computing the distribution  $p(d)$  that makes exact Bayesian inference intractable in realistic tasks.



because perceptual states approximate true posteriors under the generative model, their epistemic standing is positive in light of the model-encoded prior assumptions. Perceptual states are *justified* justifiers after all.

This epistemological picture needs three urgent clarifications. First, the inferential transitions involved in PP should not be construed in terms of a linear view of inferential justification (BonJour, 1985). Instead, PP is naturally construed as involving a coherentist element, whereby an interpretation of the input is selected that best fits into the preexisting model of the world (Gładziejewski, 2021a).

Second, perceptual inferences are constrained by raw sensory input. The input embodies statistical patterns that are a function of the causal structure of the world. At the same time, these patterns serve as a tribunal against which generative models are tested. Because of the way that the structure of the generative model depends on the structure of the sensory signal, the brain should develop generative models that represent the *most likely* causes the sensory states. Inference in PP is thus co-determined by both coherence-maximization and the ability to predict model-independent sensory patterns. The PP-based view of perceptual justification is best cast as a form of “foundherentism” (see Gładziejewski, 2021a).

Third, according to PP, perceptual justification is conditional on the rational standing of the perceptual priors. This raises a worry: aren’t *priors* acting as unjustified justifiers in this story (see Ghijssen, 2021)? I think that the answer is “no” (Gładziejewski, 2021a). On the PP approach, priors are either learned from raw sensory data or adjusted in response to sensory data. This learning is based on iterated Bayesian inference. As such, the priors have a rational etiology of their own. Furthermore, the idea that perception provides *unconditional* justification can be cashed out in terms of epistemic convergence. Perceivers usually converge on certain priors through perceptual learning. From this perspective, perceptual states can be unconditionally justified to the degree to which they are inferred from priors on which Bayesian perceptual learning convergences (Gładziejewski, 2017, 2021a).

### **3.2. Phenomenal force in Bayesian brains**

My main point is that the preceding story about perceptual justification can account for the rational contribution that the PF of visual experience makes to cognition. Specifically, within the PP framework, each of the three constituents of PF has been proposed to correspond to a hypothesis generated through unconscious (approximate) Bayesian inference. The claim I am making is conditional. If the PP-based accounts of the constituents of PF turn out true, it would mean that it is in virtue of having a rational etiology that the PF bestows the perceptual states

with the ability to justify beliefs. Below I explain how the Bayesian view applies to each of the constituents of PF.

### 3.2.1. Perceptual objectification as inference

Perceptual objectification, as introduced in Section 2, is tightly connected to a process of binding sensory features into representations of physical objects. Note that in PP, perception is a matter of forming a model of the causal structure that produces the pattern registered by the sensory apparatus. Within this framework, sensory binding is inferential. Object perception corresponds to a case where a *common cause* is inferred to account for a sensory pattern (Hohwy, 2013; Wiese, 2018; for related proposals, see e.g. Körding et al., 2007; Parise, Spence, & Ernst, 2012).

Take the following simple example. Suppose that a visual system produces a low-level representation of sensory features that participate in a stable regularity. Say, there is a sequence of edges instantiated in a location within one's visual field. The edges are co-linearly organized to form a shape. In the same location in the visual field, a roughly uniform patch of color is instantiated. Now, we may imagine that multiple perceptual hypotheses compete to explain the sensory pattern. One such candidate hypothesis treats parts of the pattern as produced and sustained by a common worldly cause (say, a red circular object). The idea is that the generative model stores assumptions that enable it to inferentially decide between this common-cause hypothesis and its alternatives. On a natural construal, the model stores prior assumptions about how likely different causal structures are, as well as likelihood priors that predict the most likely sensory patterns under respective causal hypotheses (Körding et al., 2007). Because the reliable correlation between the parts of the pattern in our toy example fits well with a common-cause hypothesis (it is likely to occur under the hypothesis), this hypothesis constitutes the winning posterior.

This sort of treatment retains the connection between perceptual objectification and the predictive structure of vision. The generative model is, after all, predictive. For example, under a common-cause explanation of the sensory input in our toy example, the generative model may predict that manipulating the object should result in a correlated change of the bound features (for instance, the features will be spatiotemporally continuous under manipulation); or that the features are mutually statistically independent, conditioned on the common cause (for example, the position of the color patch provides no additional information about the position of the shape). A natural extension of this simple idea is to claim that once the generative model

postulates a common worldly cause to explain a sensory pattern, it predicts the constancy-invariance, completion, movement-independence, and object-interaction (see also Gładziejewski 2021b; Hohwy, 2013; Madary, 2017; Seth, 2014; Wilkinson, 2020).<sup>9</sup>

Lastly, the inferential view of objectification sheds further light on the contrast cases discussed earlier. In the case of a color uniformly filling the whole visual field, the sensory pattern is too simple for a determinate cause to be inferred. In the cases of synesthesia and eye-floaters, the sensory patterns are too impoverished to merit an explanation in terms of a worldly cause (see also Seth, 2014). In those cases, the visual system presumably opts for causally “shallow” models (Hohwy, 2013) that treat the causes underlying the pattern as situated close to the sensory periphery of the perceiver (hence, as subject-dependent). In the case of depicted objects, the causal models are presumably deeper in that they treat the cause as subject-independent. But, as previously discussed, they still lack the full predictive structure of normal object perception.

### 3.2.2. Perceptual reality monitoring as inference

On the PP view, perception relies on a top-down, imagery-like process of predicting the sensory stream. This further underscores the need to distinguish perceptual from non-perceptual activity in modality-specific areas. The relevant difference is not simply between endogenous and exogenous activity. Rather, the difference is between endogenous *perceptual* activity that is heavily exogenously constrained by the prediction error signals, and an activity unconstrained, or significantly less constrained, by such signals (as when the generative model is used for non-perceptual, off-line purposes). Thus, PP is not only consistent with, but also further justifies the positing of a perceptual reality monitoring mechanism.

A number of authors working within the PP and related (Bayesian, generative-model-based) frameworks have proposed that perceptual reality monitoring is subserved by a second-order Bayesian inference (Fleming, 2020; Fleming & Daw, 2017; Gershman, 2019; Lau, 2008, 2019). The common idea underlying these proposals is that frontal areas (presumably involving the anterior prefrontal cortex) take first-order activity in modality-specific areas (say, in the visual cortex) as input, and produce metacognitive estimates of this latter activity as output

---

<sup>9</sup> Note that this sheds additional light on the view that off-line quasi-perceptual states can exhibit the predictive structure of objectification. Generative models can endogenously simulate causal structures to simulate top-down predictions of lower-level perceptual activity. Although these simulations of first-order perception will differ from actual perception (for example, be less constrained by the prediction errors), this is a difference of degree, not of (computational, neural) kind.

posteriors. What is being estimated is the source of the first-order activity, that is, whether the activity is perceptual, non-perceptual (as in imagery), or arises through pure stochastic noise (Lau, 2019). The inference is mediated by prior assumptions regarding the first-order activity.

Here is a simple illustration of how this second-order inference may proceed. Suppose, for simplicity, that the reality monitoring system is to decide between a hypothesis that some piece of first-order activity in the visual cortex reflects the external world (that is, it is perceptual), and a hypothesis on which the same activity reflects imagery. The crux is to employ priors about what features of the first-order activity are to be expected under the respective hypotheses. Consider two plausible likelihood prior assumptions that a metacognitive inference of this sort could rely on. First, perceptual contents tend to be significantly more detailed than the contents of imagery, which is reflected by the fact that perception involves stronger neural signals than imagery (Dijkstra, Kok, & Fleming, 2021). Second, subjects tend to exert less cognitive control over perceptual contents than over imagined contents (Dijkstra, Kok, & Fleming, 2021). Hence, the respective hypotheses entail differing predictions regarding the first-order activity. To disambiguate between the hypotheses, the metacognitive engine can rely on information regarding how detailed the first-order content is and how much control the subject has over it.

Note that the considerations above still leave a fair degree of ambiguity in deciding between the perception hypothesis and the imagery hypothesis (Dijkstra, Kok, & Fleming, 2021; see also Farkas, 2013). Perceptual contents can lack detail in non-ideal conditions, while imagined contents can be sometimes fairly detailed. Furthermore, imagery states can be sometimes spontaneously generated (as in mind wandering) without much cognitive control, while perceptual contents can be voluntarily changed through shifts of attention. However, this sort of uncertainty is precisely what motivates construing perceptual metacognition as Bayesian inference. It could also explain source confusions—like mistaking perception for imagery—that occur in experimental conditions designed to exploit the ambiguities in first-order signals (Dijkstra & Fleming, 2021; Dijkstra et al., 2021; Perky, 1910).

### **3.2.3. Interoceptive appraisal as inference**

Previous considerations about the interoceptive appraisal side of the perceptual sense of reality diverge from a default “depictive” view of PF. It is common to assume that the sense of reality is attached to the perceptual representation of what objects comprise the visual scene and what properties they possess. But on the present view, the interoceptive side of PF is geared towards tracking the hedonic or motivational value. Thus, to connect this evaluative side of the

perceptual sense of reality to the literature on PP, we need an account of how the prediction error minimization allows the cognitive system to estimate the hedonic or motivational value.

Two ideas that have emerged in the literature are of relevance here. First, according to PP, interoception—just like exteroception—is a matter of causal inference. The brain takes the streams of interoceptive signals as input and produces, through the minimization of interoceptive prediction error, posteriors that explain the causes of those signals (Allen & Tsakiris, 2019; Seth, 2013). For example, emotional states may correspond to hypotheses that aim to causally explain the interoceptive stream (is my stomach contacting because I am afraid, disgusted, or falling in love?).

Second, the notion of interoceptive inference has been central in PP-based accounts of how the brain's representation of self is constructed (Allen & Tsakiris, 2019; Hohwy & Michael, 2017; Letheby & Gerrans, 2017). Roughly, the idea is that the basic sense of self emerges through a common cause inference. A singular, persisting entity, the self, is posited to explain the reliable correlations in interoceptive signals. One such reliable interoceptive pattern that gives rise to the brain's self-model is homeostatic. The self, as posited by the self-model, is a thing that tends to undergo homeostatically viable interoceptive states, that is states with positive hedonic valence. In accordance with PP, this self-model structures the top-down interoceptive predictions. Thus, actively seeking states of affairs that bring about hedonically positive interoceptive responses is equivalent to seeking sensory evidence that confirms the brain's self-model.

This process of using the self-model to interpret the incoming signals, as well as to seek evidence that favors the model, is thought to underpin the experience of being affectively connected to the world (Gerrans, 2019; Deane, Miller, & Wilkinson, 2020). As such, the interoceptive background state discussed in Section 2.3 is generated through inference. This idea allows us to understand the disturbance of the background interoceptive state in depersonalization/derealization. In those cases, due to de-affectualization, the brain is stuck in a state in which it gets persistently flattened affective (interoceptive) feedback from the world. This means that the brain obtains interoceptive signals which disconfirm at least some prior assumptions about the self. This is manifested phenomenologically as the self-estrangement that characterizes depersonalization (Gerrans, 2019; Deane, Miller, & Wilkinson, 2020). But the same disturbance has an exteroceptive side in that the world fails to feel affectively salient and meaningful, which corresponds to derealization (Gerrans, 2019; Deane, Miller, & Wilkinson, 2020).

#### 4. Is subpersonal etiology epistemically relevant?

In this remaining section, I address three possible concerns about the present proposal.

##### 4.1. Why treat the etiology of PF as epistemically relevant?

Why shouldn't we settle on a dogmatist position which takes the epistemic role of PF at face value? Why claim at all that the source of the epistemic import of PF lies in its etiology? There are, I think, two rationales that speak in favor of the etiology-matters approach.

Consider, first, the analogy to the "autonoetic" PF that characterizes the experience of reliving events in episodic memory. When we mentally travel in time to our past, we (usually) have a strong sense that the episode we remember actually happened the way we remember it. However, empirical research has established that episodic memory is a constructive process in which general knowledge is heavily deployed to fill in fragmentary memory traces (cf. De Brigard, 2014). Of particular relevance here, there is a reality monitoring mechanism at play in episodic memory that gives rise to the very sense that a mental event is an authentic recollection (Johnson, Hashtroudi, & Lindsay, 1993). Occasionally, a state that feels like mental time travel may be in fact an imagery state misidentified as a memory. Because of its constructive nature, episodic memory is considerably more error-prone than is readily apparent from the first-person perspective.

Say, now, that I experience a vivid childhood memory of getting lost in a shopping mall. Pretheoretically, it may seem that my strong sense that the event in question actually happened justifies me in believing that I got lost in a shopping mall as a child. However, given facts about how episodic memories are produced, it seems reasonable to think that this folk-epistemological view requires adjustments. In particular, the justification I obtain for my belief based on my mnemonic experience is *weaker* than it pretheoretically appears. Regardless of how the pretheoretical view is to be adjusted exactly, the very claim that the epistemic import of autonoetic PF depends on its etiology seems plausible. At the very least, it does not seem less plausible than the claim that we should ignore etiology altogether. If this holds in the case of the PF of memory, then the same should presumably apply to the PF of visual perception.

The second rationale for the etiology-matters view is related to considerations of explanatory depth. If theory  $T_1$  treats a phenomenon  $P$  as an unexplained datum or a postulate, while a rival theory  $T_2$  provides a principled explanation for  $P$ , then, all else being equal,  $T_2$  is preferable to  $T_1$  due to being explanatorily deeper. My point is that the inferentialist view of PF

is preferable over the dogmatist view on account of its explanatory depth. As far as I can tell, dogmatism fails to deliver an explanation of what it is about PF that enables the perceptual states that have it to confer justification on other mental states. Dogmatists simply follow common sense in positing that this is how things work. In contrast, the inferentialist view explains where the justificational role of PF comes from. Under an assumption that all else is indeed equal, the etiology-matters account is preferable to the dogmatist one.

#### **4.2. Can subpersonal Bayesian inference generate justification?**

A critic might argue that the discussion above misses the point of denying that the origin of PF has epistemic import. The crux of the problem lies not in invoking the etiology as such but in the fact the etiology to which the present account appeals is *of the wrong kind*. The account on offer posits that the epistemically relevant etiology involves subpersonal states and processes. But the only level at which epistemic value is to be found, the critic may argue, is the personal level of states and processes properly attributable to people.

A systematic defense of the position that subpersonal phenomena are epistemically appraisable and capable of justification conferral is beyond the scope of the present article. I can only briefly sketch out some reasons that favor such a view. A recently proposed argument is based on an investigation of epistemically relevant cognitive mechanisms, that is, mechanisms that underlie rational (evidence-responsive, coherence-maximizing) belief acquisition and maintenance (Carter & Rupert, 2021). It turns out that these mechanisms are primarily subpersonal (Carter & Rupert, 2021). The upshot, roughly, is that we need to treat the subpersonal level as having epistemic import in order to give justice to the full scope of epistemically relevant cognition. This generalizes the conclusion made above: allowing the subpersonal level as epistemically relevant greatly expands the explanatory depth of epistemology.

An alternative argumentative strategy consists in arguing *against* the position that epistemic value is exclusively found at the personal level. A proponent of this view presumably holds that personal-level states or processes are such that: (1) they have certain features whose possession is necessary for a cognitive state to be epistemically relevant (epistemically appraisable and capable of justification transfer), and (2) these features are uniquely found at the personal level. But what are those features? When we take a close look at the obvious candidates, they either turn out not to be necessary for epistemic import or not confined to the personal level.

One such candidate feature might be conscious accessibility to the subject. However, it has been repeatedly pointed out that unconscious beliefs and inferences can sometimes confer justification (cf. Siegel, 2017; Jenkin, 2020). For example, a belief formed through inattentive reasoning that escaped the subject's awareness can still be justified in light of this reasoning. Conscious accessibility is not a necessary condition on having epistemic import.

Voluntary control is another seemingly plausible candidate. Personal-level states and processes are ones over which the subject enjoys, or so it seems, a fair degree of control. However, it seems that a purely spontaneous chain of reasoning can be epistemically appraisable even if neither its initiation nor the transitions that lead from premises to a conclusion are under the voluntary control of the subject (Jenkin, 2020). Voluntary control is not a requirement for having epistemic import.

Lastly, the key to epistemic import might lie in being answerable to reasons. For example, beliefs are revisable in response to counterarguments. However, under a relaxed but nontrivial interpretation of this feature, some subpersonal states are answerable to reasons. Of particular interest here is research that establishes that the priors that, according to the Bayesian view, participate in perceptual inferences can be revised when the subject is exposed to sensory stimulation that conflicts with them (for philosophically oriented discussions of this work, see Gładziejewski 2021b; Rescorla, 2018). Answerability to reasons does not look like a distinctive feature of personal-level mentality. Overall, the idea that epistemic value can be found at the subpersonal level is far from hopeless.

#### **4.3. Is this view committed to externalism about epistemic justification?**

Yet another possible worry about the present proposal pertains to more general assumptions regarding the nature of epistemic justification. Dogmatist accounts of the epistemic role of PF naturally align with internalism about epistemic justification. It might be objected that I tacitly changed the rules of the game by presupposing an externalist view of epistemic justification.

I do not think that this diagnosis is quite right. My argument in the present paper does *not* require subscribing to externalism about justification. The view on offer here is in line with the position that PF's epistemic import is grounded in rational relations between states that are internal to the cognitive life of the subject (see Gładziejewski 2021a). This is the "mentalist" variant of internalism (Conee & Feldman, 2001). So the present proposal is compatible with at least one, permissive form of epistemic internalism.



## 5. Conclusions

The purpose of the present paper was to establish an account of the phenomenal force (PF) of visual perceptual experience and its rational involvement in cognition. I showed how distinct aspects or constituents of PF could be rooted in unconscious probabilistic inference. On the view introduced here, which stands in stark contrast to existing foundationalist accounts, it is through having a rational etiology that PF contributes to perceptual justification.

Two caveats are in order. First, the discussion in this paper focused on the PF of visual perception. I leave open the question of whether the scope of the present position can be extended to shed light on the epistemology of non-visual modalities, or on the epistemology of purely intellectual intuitions that have also been construed as possessing PF (Chudnoff, 2013). Second, because of how heavily the present position relies on work in cognitive science, its truth cannot be established on a priori grounds alone. It lives or dies depending on whether future research in cognitive science reinforces the idea that the sense of reality has a rational etiology.

## ACKNOWLEDGEMENTS

I would like to thank Michał Piekarski for encouraging conversations. Work on this paper was supported by National Science Center in Poland (grant number, 2019/33/B/HS1/00677).

## REFERENCES

- Allen, M., & Tsakiris, M. (2019). The body as first prior: Interoceptive predictive processing and the primacy of self-models. In M. Tsakiris & H. De Preester (Eds.). *The interoceptive mind: From homeostasis to awareness* (pp. 27–45). Oxford: Oxford University Press.
- Bogacz, R. (2017). A tutorial on the free-energy framework for modelling perception and learning. *Journal of Mathematical Psychology*, 76, B, 198–211.
- BonJour, L. (1985). *The structure of empirical knowledge*. Cambridge, MA: Harvard University Press.
- Burge, T. (2010). *Origins of objectivity*. Oxford: Oxford University Press.
- Carter, J. D., & Rupert, R. (2021). Epistemic value in the subpersonal vale. *Synthese*, 198, 9243–9272.

- Chudnoff, E. (2013). *Intuition*. Oxford: Oxford University Press.
- Church, J. (2013). *Possibilities of perception*. Oxford: Oxford University Press.
- Clark, A. (2012). Dreaming the whole cat: Generative models, predictive processing, and the enactivist conception of perceptual experience. *Mind*, *121*, 753–771.
- Colombetti, G., & Ratcliffe, M. (2012). Bodily feeling in depersonalization: A phenomenological account. *Emotion Review*, *4*, 145–150.
- Conee, E., & Feldman, R. (2001). Internalism defended. *American Philosophical Quarterly*, *38*, 1–18.
- Damasio, A. (1994). *Descartes' error: Emotion, reason and the human brain*. New York, NY: Putnam.
- Deane, G., Miller, M., & Wilkinson, S. (2020). Losing ourselves: Active inference, depersonalization, and meditation. *Frontiers in Psychology*, *11*, 539726.
- De Brigard, F. (2014). Is memory for remembering? Recollection as a form episodic hypothetical thinking. *Synthese*, *191*, 155–185.
- Dijkstra, N., & Fleming, S. (2021). Fundamental constraints on distinguishing reality from imagination. PsyArXiv preprint: <https://psyarxiv.com/bw872>
- Dijkstra, N., Kok, P., & Fleming, S. (2021). Perceptual reality monitoring: Neural mechanisms dissociating imagination from reality. PsyArXiv preprint: <https://psyarxiv.com/zngeq>
- Dijkstra, N., Mazor, M., Kok, & P. Fleming, S. (2021). Mistaking imagination for reality: congruent mental imagery leads to more liberal perceptual detection. *Cognition*, *212*, 1-9.
- Dokic, J., Martin, J-R. (2017). Felt reality and the opacity of perception. *Topoi*, *36*, 299–309.
- Farkas, K. (2013). A sense of reality. In F. Macpherson & D. Platchias (Eds.). *Hallucination* (pp. 399–414). Cambridge, MA: The MIT Press.
- Fleming, S. (2020). Awareness as inference in a higher order state space. *Neuroscience of Consciousness*, *6*, niz020.
- Fleming, S., & Daw, N. D. (2017). Self-evaluation of decision-making: a general Bayesian framework for metacognitive computation. *Psychological Review*, *124*, 91–114.
- Friston, K. J. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B*, *360*, 815–836.
- Gerrans, P. (2019). Depersonalization disorder, affective processing and predictive coding. *Review of Philosophy and Psychology*, *10*, 401–418.

Gershman, S.J. (2019). The generative adversarial brain. *Frontiers in Artificial Intelligence*, 2, Article 18.

Ghijzen, H. (2014). Phenomenalist dogmatist experientialism and the distinctiveness problem. *Synthese*, 191, 1549–1566.

Ghijzen, H. (2021). Predictive processing and foundationalism about perception. *Synthese*, 198, 1751–1769 .

Gładziejewski, P. (2017). Evidence of the senses: A predictive processing-based take on the Sellarsian dilemma. In T. Metzinger & W. Wiese (Eds). *Philosophy and predictive processing*. MIND Group. Available at: <https://predictive-mind.net/papers/the-evidence-of-the-senses>.

Gładziejewski, P. (2021a). Perceptual justification in the Bayesian brain: A foundherentist account. *Synthese*, 199, 11397–11421.

Gładziejewski, P. (2021b). Un-debunking ordinary objects with the help of predictive processing. *British Journal for the Philosophy of Science*, doi: <https://doi.org/10.1086/715105>.

Hohwy, J. (2020). New directions in predictive processing. *Mind & Language*, 35, 209–223.

Hohwy, J., & Michael, J. (2017). Why should any body have a self? In F. de Vignemont & A. Alsmith (Eds.). *The subject's matter: Self-consciousness and the body* (pp. 363–392). Cambridge, MA: The MIT Press.

Huemer, M. (2001). *Skepticism and the veil of perception*. Lanham, MD: Rowman and Littlefield Publishers Inc.

Jenkin, Z. (2020). The epistemic role of core cognition. *The Philosophical Review*, 129, 251–298.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114, 3–28.

Kiefer, A. (2017). Literal perceptual inference. In T. Metzinger, & W. Wiese (Eds.). *Philosophy and predictive processing*. MIND Group. Available at: <https://predictive-mind.net/papers/literal-perceptual-inference>.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLOS One*, 9, e943.

Lau, H. (2008). A higher order Bayesian decision theory of consciousness. *Progress in Brain Research*, 168, 35-48.

Lau, H. (2019). Consciousness, metacognition, & perceptual reality monitoring. PsyArXiv preprint: <https://psyarxiv.com/ckbyf/>

- Letheby, C., & Gerrans, P. (2017). Self unbound: Ego dissolution in psychedelic experience. *Neuroscience of Consciousness*, *1*, 1–11.
- Madary, M. (2017). *Visual phenomenology*. Cambridge, MA: The MIT Press.
- Matthen, M. (2010). Two visual systems and the feeling of presence. In N. Gangopadhyay, M. Madary & F. Spicer (Eds.). *Perception, action, and consciousness: Sensorimotor dynamics and two visual systems*. Oxford: Oxford University Press.
- Medford, N. (2012). Emotion and the unreal self: Depersonalization disorder and de-affectualization. *Emotion Review*, *4*, 139–144.
- Munton, J. (2018). The eye's mind: Perceptual process and epistemic norms. *Philosophical Perspectives*, *31*, 317–347.
- Noë, A. (2002). On what we see. *Pacific Philosophical Quarterly*, *83*, 57–80.
- Parise, C. V., Spence, C., & Ernst, M. O. (2012). When correlation implies causation in multisensory integration. *Current Biology*, *22*, 46–49.
- Pearson, J. (2019). The human imagination: the cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience*, *20*, 624–634.
- Perky, C. W. (1910). An experimental study of imagination. *American Journal of Psychology*, *21*, 422–452.
- Prinz, J. (2006). *Gut reactions: A perceptual theory of emotions*. Oxford: Oxford University Press.
- Pryor, J. (2000). The skeptic and the dogmatist. *Noûs*, *34*, 517–549.
- Rao, R. P. N., & Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–87.
- Ratcliffe, M. (2008). *Feelings of being: Phenomenology, psychiatry and the sense of reality*. Oxford: Oxford University Press.
- Rescorla, M. (2018). An interventionist approach to psychological explanation. *Synthese*, *195*, 1909–1940.
- Sass, L. A., & Parnas, J. (2003). Schizophrenia, consciousness, and the self. *Schizophrenia Bulletin*, *29*, 427–444.
- Sass, L., Pienkos, E., Nelson, B., & Medford, N. (2013). Anomalous self-experience in depersonalization and schizophrenia: A comparative investigation. *Consciousness and Cognition*, *22*, 430–441.
- Sellars, W. (1956). Empiricism and the philosophy of mind. In: H. Feigl & N. Scriven (Eds.). *Minnesota studies in the philosophy of science* (pp. 253–329). Minneapolis, MN: University of Minnesota Press.

Seth, A. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in the Cognitive Sciences*, *17*, 565–573.

Seth, A. (2014). A predictive processing theory of sensorimotor contingencies: Explaining the puzzle of perceptual presence and its absence in synesthesia. *Cognitive Neuroscience*, *5*, 97–118.

Siegel, S. (2010). *The Contents of Visual Experience*. New York, NY: Oxford University Press.

Siegel, S. (2017). *The Rationality of Perception*. Oxford: Oxford University Press.

Siegel, S., Silins, N. (2015). The epistemology of perception. In M. Matthen (Ed.). *Oxford handbook of philosophy of perception*. New York, NY: Oxford University Press.

Simons, J. S., Garrison, J. R., & Johnson, M. K. (2017). Brain mechanisms of reality monitoring. *Trends in Cognitive Sciences*, *21*, 462–473.

Skrzypulec, B. (2018). Thisness and visual objects. *Review of Philosophy and Psychology*, *9*, 17–32.

Teng, L. (2016). Is phenomenal force sufficient for immediate perceptual justification? *Synthese*, *195*, 637–656.

Tucker, C. (2010). Why open-minded people should endorse dogmatism. *Philosophical Perspectives*, *24*, 529–545.

Vance, J. (2015). Cognitive penetration and the tribunal of experience. *Review of Philosophy and Psychology*, *6*, 641–663.

Vernazzani, A. (2021). Do we see facts? *Mind & Language*, doi:10.1111/mila.12336.

Wiese, W. (2018). *Experienced wholeness: Integrating insights from Gestalt theory, cognitive neuroscience, and predictive processing*. Cambridge, MA: The MIT Press.

Wilkinson, S. (2020). Distinguishing volumetric content from perceptual presence within a predictive processing framework. *Phenomenology and the Cognitive Sciences*, *19*, 791–800.

Windt, J. (2015). *Dreaming: A conceptual framework for philosophy of mind and empirical research*. Cambridge, MA: The MIT Press.

Zeman, A., Milton, F., Della Sala, S., Dewar, M., Frayling, T., Gaddum, J., Hattersley, A., Heuerman-Williamson, B., Jones, K., MacKisack, M., & Winlove, C. (2020). Phantasia—The psychological significance of lifelong visual imagery vividness extremes. *Cortex*, *130*, 426–440.